

Jarrell, Noble

182151

From: Ramirez, Delia
Sent: Tuesday, March 14, 2006 1:44 PM
To: Jarrell, Noble
Subject: 10/673786

Hi,

I would like to request the following search:

1. SEQ ID NO:1-2 in the nucleic acid databases (commercial and interference)
2. SEQ ID NO:2 in the protein databases (commercial and interference)

Thank you very much,

Delia M. Ramirez, Ph.D.
Patent Examiner
Recombinant Enzymes-Art Unit 1652
USPTO
400 Dulany Street, Remsen Bldg., 2D74, Mail room 2C70
Alexandria, VA 22314
(571) 272-0938
delia.ramirez@uspto.gov

Noble
SPR Fin 3/17/06
SONL
2NA Compaugnan
1AA

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2006, 14:55:19 ; Search time 2975 Seconds
(without alignments)
3310.531 Million cell updates/sec

Title: US-10-673-786A-1
Perfect score: 1191
Sequence: 1 atgtttgagaacattaccgc.....cgattgtgcagtgctgttaa 1191

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA_Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1191	100.0	1191	6	US-10-369-493-24472	Sequence 24472, A
2	1191	100.0	1191	7	US-10-673-786A-1	Sequence 1, Appli
3	1191	100.0	14759	6	US-10-893-671-87	Sequence 87, Appl
4	655.8	55.1	1188	6	US-10-369-493-44812	Sequence 44812, A
5	510	42.8	966	6	US-10-369-493-23918	Sequence 23918, A
6	496.2	41.7	2133	9	US-10-450-763-23995	Sequence 23995, A
7	486.2	41.7	3222	9	US-10-450-763-25886	Sequence 25886, A
8	493.4	41.4	7977	9	US-10-795-159-536	Sequence 536, App
9	493.4	41.4	908766	9	US-10-795-159-685	Sequence 685, App
10	471.8	39.6	1191	7	US-10-275-026A-187	Sequence 187, App
11	466.8	39.2	17381	9	US-10-915-740A-40	Sequence 40, Appl
12	466.8	39.2	2242716	9	US-10-915-740A-1068	Sequence 1068, Ap
13	458.2	38.5	1830121	7	US-10-329-670-1	Sequence 1, Appli
14	458.2	38.5	1830121	8	US-10-158-865-1	Sequence 1, Appli
15	458.2	38.5	1830121	9	US-10-981-687-1	Sequence 1, Appli
16	363	30.5	2220	6	US-10-450-763-17608	Sequence 17608, A
17	331.8	27.9	1185	6	US-10-369-493-35511	Sequence 35511, A
18	327.8	27.5	1191	6	US-10-369-493-32111	Sequence 32111, A
19	320.8	26.9	1188	6	US-10-369-493-31184	Sequence 31184, A
20	320.8	26.9	1194	6	US-10-369-493-28425	Sequence 28425, A
21	310.6	26.1	1194	6	US-10-369-493-37712	Sequence 37712, A
22	306.6	25.7	24417	6	US-10-216-209-1	Sequence 1, Appli
23	297.4	25.0	1093	9	US-10-450-763-13154	Sequence 13154, A

24	288.8	24.2	1188	6	US-10-369-493-32346	Sequence 32346, A
25	287.6	24.1	1185	6	US-10-369-493-37410	Sequence 37410, A
26	285.2	23.9	1191	6	US-10-369-493-39515	Sequence 39515, A
27	285.2	23.9	1191	6	US-10-369-493-39895	Sequence 39895, A
28	285.2	23.9	1194	6	US-10-369-493-39147	Sequence 39147, A
29	284	23.8	1188	6	US-10-369-493-28514	Sequence 28514, A
30	284	23.8	1188	6	US-10-369-493-31273	Sequence 31273, A
31	274.8	23.1	100848	7	US-10-672-787-39	Sequence 39, Appl
32	267.8	22.5	35133	9	US-10-915-740A-65	Sequence 65, Appl
33	262.8	22.1	1188	6	US-10-369-493-44841	Sequence 44841, A
34	257	21.6	1194	6	US-10-369-493-47344	Sequence 47344, A
35	255.4	21.4	1194	7	US-10-422-366-31	Sequence 31, Appl
36	255.4	21.4	1194	9	US-10-925-216-31	Sequence 31, Appl
37	255.4	21.4	1194	9	US-10-903-582-31	Sequence 31, Appl
38	255.4	21.4	1194	9	US-10-918-401A-31	Sequence 31, Appl
39	255.4	21.4	1194	9	US-10-969-245-31	Sequence 31, Appl
40	250.2	21.0	1185	6	US-10-369-493-33145	Sequence 33145, A
41	248.6	20.9	1185	6	US-10-369-493-32909	Sequence 32909, A
42	244.8	20.6	1456	7	US-10-425-114-22011	Sequence 22011, A
43	243.4	20.4	1203	6	US-10-369-493-41174	Sequence 41174, A
44	243.4	20.4	2731748	7	US-10-297-465A-1	Sequence 1, Appli
45	241.6	20.3	1260	7	US-10-422-366-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-10-369-493-24472
; Sequence 24472, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24472
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-369-493-24472

Query Match	100.0%	Score 1191;	DB 6;	Length 1191;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1191;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGTTTGAGAACATTACCGCGCTCTGCGCGACCCCGATTCTGGGCCCTGGCGGATCTGTTT	60	
Db	1	ATGTTTGAGAACATTACCGCGCTCTGCGCGACCCCGATTCTGGGCCCTGGCGGATCTGTTT	60	
Qy	61	CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGACG	120	
Db	61	CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGACG	120	
Qy	121	GGCAAAACCCGGTACTGACACGGTGAAAAGCTGAACAGTATCTGCTCGAAAATGAA	180	
Db	121	GGCAAAACCCGGTACTGACACGGTGAAAAGCTGAACAGTATCTGCTCGAAAATGAA	180	
Qy	181	ACCACCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGACG	240	
Db	181	ACCACCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGACG	240	
Qy	241	CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGTCGTCGACGGCAGACT	300	

```
Db 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTGCTGCGACGGCACAGACT 300
Qy 301 CCGGGGGGCACTGGCGCACTACGGGTGGCTGCGGATTTCTTGGCAAAAATACAGGTTT 360
Db 301 CCGGGGGGCACTGGCGCACTACGGGTGGCTGCGGATTTCTTGGCAAAAATACAGGTTT 360
Qy 361 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCGCAACCATGAAGCGTCTTTAACTCTGCA 420
Db 361 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCGCAACCATGAAGCGTCTTTAACTCTGCA 420
Qy 421 GGTCTGGAAGTTCTGTAATAGCTTTATTATGATCGGAAAATCACACTCTTGACTTCGAT 480
Db 421 GGTCTGGAAGTTCTGTAATAGCTTTATTATGATCGGAAAATCACACTCTTGACTTCGAT 480
Qy 481 GCACCTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTCTTCCATGGCTGC 540
Db 481 GCACCTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTCTTCCATGGCTGC 540
Qy 541 TGGCATTAACCAACCGGTATCGACCCCTACGCTGGAACAATGGCAAAACACTGGCACAACTC 600
Db 541 TGGCATTAACCAACCGGTATCGACCCCTACGCTGGAACAATGGCAAAACACTGGCACAACTC 600
Qy 601 TCCGTTGAGAAAGCTGTTTACCCTGTTTGAATTCGCTTACCAGGGTCTTTGCCCGTGGT 660
Db 601 TCCGTTGAGAAAGCTGTTTACCCTGTTTGAATTCGCTTACCAGGGTCTTTGCCCGTGGT 660
Qy 661 CTGGAAGAAGTGTGAAGGACTCGCGCTTTTCCGGCTATGCAATAAAGAGCTGATTTGTT 720
Db 661 CTGGAAGAAGTGTGAAGGACTCGCGCTTTTCCGGCTATGCAATAAAGAGCTGATTTGTT 720
Qy 721 GCCAGTTCCTACTCTAAACCTTTGGGCTGTGACAAAGAGCGTGTGGCGCTTGACTCTG 780
Db 721 GCCAGTTCCTACTCTAAACCTTTGGGCTGTGACAAAGAGCGTGTGGCGCTTGACTCTG 780
Qy 781 GTTCTGCGGACAGTGAACCGTTGATCGCGCAATTCAGCCAAATGAAGCGGCGATTCGC 840
Db 781 GTTCTGCGGACAGTGAACCGTTGATCGCGCAATTCAGCCAAATGAAGCGGCGATTCGC 840
Qy 841 GCTAACTACTCTAAACCCACAGCACACGGCGCTTCTGTTGTCACCATCTCTGAGCAAC 900
Db 841 GCTAACTACTCTAAACCCACAGCACACGGCGCTTCTGTTGTCACCATCTCTGAGCAAC 900
Qy 901 GATCGGTTAGTGGGATTTGGGAAACAAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960
Db 901 GATCGGTTAGTGGGATTTGGGAAACAAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960
Qy 961 ATGCGTCAGTTGTTTCTCAATAGCTGCGGAAAGCGCAACCGCGACTTCAGCTTT 1020
Db 961 ATGCGTCAGTTGTTTCTCAATAGCTGCGGAAAGCGCAACCGCGACTTCAGCTTT 1020
Qy 1021 ATCATCAAAAGAGAGCGGATGTTCTCTTCACTGCGGCTGACAAAGAAACAAAGTCTGCGT 1080
Db 1021 ATCATCAAAAGAGAGCGGATGTTCTCTTCACTGCGGCTGACAAAGAAACAAAGTCTGCGT 1080
Qy 1081 CTGCGGAGAGTTTGGGATATAGCGGTGCTTCTGCTGCGTAAATGTTGGCGGGATG 1140
Db 1081 CTGCGGAGAGTTTGGGATATAGCGGTGCTTCTGCTGCGTAAATGTTGGCGGGATG 1140
Qy 1141 ACACAGATACATGGCTCGCTGCGGAGCGATTCGCGAGTCTGTAA 1191
Db 1141 ACACAGATACATGGCTCGCTGCGGAGCGATTCGCGAGTCTGTAA 1191
```

RESULT 2

```
US-10-673-786a-1
; Sequence 1, Application US/10673786A
; Publication No. US20040132165A1
; GENERAL INFORMATION:
; APPLICANT: AKHVERDIAN, VALERY ZAVENOVICH
; APPLICANT: SAVRASOVA, EKATERINA ALEKSEEVNA
; APPLICANT: KAPLAN, ALLA MARKOVNA
; APPLICANT: LOBANOV, ANDREY OLEGOVICH
; APPLICANT: KOZLOV, YURI IVANOVICH
```

```
; TITLE OF INVENTION: METHOD FOR PRODUCING L-THREONINE USING BACTERIA
; FILE OF INVENTION: BELONGING TO THE GENUS ESCHERICHIA
; CURRENT APPLICATION NUMBER: US/10/673,786A
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: PCT/JP03/02067
; PRIOR FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: 2002104983
; PRIOR FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1191)
US-10-673-786a-1
```

```
Query Match 100.0%; Score 1191; DB 7; Length 1191;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTTGGAGAACATTACCGCGCTCTGCGACCCCGATTCTGGGCTTGGCCGATCTGTTT 60
Db 1 ATGTTTGGAGAACATTACCGCGCTCTGCGACCCCGATTCTGGGCTTGGCCGATCTGTTT 60
Qy 61 CGTGCGGATGAACGCTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 120
Db 61 CGTGCGGATGAACGCTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 120
Qy 121 GGCAAAACCCCGGTACTGACAGCGTGAAGAGCTGAACAGTATCTGCTCGAAAATGAA 180
Db 121 GGCAAAACCCCGGTACTGACAGCGTGAAGAGCTGAACAGTATCTGCTCGAAAATGAA 180
Qy 181 ACCACAAAATTTACCTCGGCAATGACGCGATCCCTGAAATTTGGTCTGCTCAGGAA 240
Db 181 ACCACAAAATTTACCTCGGCAATGACGCGATCCCTGAAATTTGGTCTGCTCAGGAA 240
Qy 241 CTGCTCTTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTCTCGCAGGCAAGACT 300
Db 241 CTGCTCTTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTCTCGCAGGCAAGACT 300
Qy 301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAATACAGGCTT 360
Db 301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAATACAGGCTT 360
Qy 361 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCGCAACCATGAAGCGTCTTTAACTCTGCA 420
Db 361 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCGCAACCATGAAGCGTCTTTAACTCTGCA 420
Qy 421 GGTCTGGAAGTTCTGTAATAGCTTTATTATGATCGGAAAATCACACTCTTGACTTCGAT 480
Db 421 GGTCTGGAAGTTCTGTAATAGCTTTATTATGATCGGAAAATCACACTCTTGACTTCGAT 480
Qy 481 GCACCTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTCTTCCATGGCTGC 540
Db 481 GCACCTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTCTTCCATGGCTGC 540
Qy 541 TGGCATTAACCAACCGGTATCGACCCCTACGCTGGAACAATGGCAAAACACTGGCACAACTC 600
Db 541 TGGCATTAACCAACCGGTATCGACCCCTACGCTGGAACAATGGCAAAACACTGGCACAACTC 600
Qy 601 TCCGTTGAGAAAGCTGTTTACCCTGTTTGAATTCGCTTACCAGGGTCTTTGCCCGTGGT 660
Db 601 TCCGTTGAGAAAGCTGTTTACCCTGTTTGAATTCGCTTACCAGGGTCTTTGCCCGTGGT 660
Qy 661 CTGGAAGAAGTGTGAAGGACTCGCGCTTTTCCGGCTATGCAATAAAGAGCTGATTTGTT 720
Db 661 CTGGAAGAAGTGTGAAGGACTCGCGCTTTTCCGGCTATGCAATAAAGAGCTGATTTGTT 720
Qy 721 GCCAGTTCCTACTCTAAACCTTTGGGCTGTGACAAAGAGCGTGTGGCGCTTGACTCTG 780
```



```

Db 721 GCCAGTTCTTCTAATAAATTTGGCTGTACAAAGAGGCTGTGGCGTTGTACTCTG 780
Qy 781 GTTGCTCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGGATTCGC 840
Db 781 GTTGCTCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGGATTCGC 840
Qy 841 GCTAACTACTCTAACCCACAGCAGACACGCGCTTCTGTTGTCGCCACCATCTCAGCAAC 900
Db 841 GCTAACTACTCTAACCCACAGCAGACACGCGCTTCTGTTGTCGCCACCATCTCAGCAAC 900
Qy 901 GATCGTTAGTGGATTTGGGAACAAGAGTGAATGAGCCAGAGCTATTCAGCGT 960
Db 901 GATCGTTAGTGGATTTGGGAACAAGAGTGAATGAGCCAGAGCTATTCAGCGT 960
Qy 961 ATGCGTCAGTTGTCGTCATAGCTGACGAGAAAGCGCAACCGCGCTTCTGTTGTCGCCACCATCTCAGCTTT 1020
Db 961 ATGCGTCAGTTGTCGTCATAGCTGACGAGAAAGCGCAACCGCGCTTCTGTTGTCGCCACCATCTCAGCTTT 1020
Qy 1021 ATCATCAACAGAACGCGCATGTTCTCCTTCAGTGGCTGACAAAGAACCAAGTCTGCGT 1080
Db 1021 ATCATCAACAGAACGCGCATGTTCTCCTTCAGTGGCTGACAAAGAACCAAGTCTGCGT 1080
Qy 1081 CTGCGGAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTCGCGGATG 1140
Db 1081 CTGCGGAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTCGCGGATG 1140
Qy 1141 ACACCAAGATACATGCTGCGTGTGCGAAGCGATTTGTCAGTGTGCTGTA 1191
Db 1141 ACACCAAGATACATGCTGCGTGTGCGAAGCGATTTGTCAGTGTGCTGTA 1191

```

RESULT 3

```

US-10-893-671-87/c
; Sequence 87, Application US/10893671
; Publication No. US20050064527A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Stuart, et. al.
; TITLE OF INVENTION: NMR COMPOSITIONS AND THEIR METHODS OF USE
; FILE REFERENCE: PKZ-043
; CURRENT APPLICATION NUMBER: US/10/893,671
; CURRENT FILING DATE: 2004-07-15
; PRIOR APPLICATION NUMBER: US/09/801,563
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/188,362
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 87
; LENGTH: 14759
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-893-671-87

```

```

Query Match 100.0%; Score 1191; DB 9; Length 14759;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTTGAGAACATTAACCGCGCTCTGCGACCCGATTCGGGCTGCGCGATCTGTTT 60
Db 2755 ATGTTTGAGAACATTAACCGCGCTCTGCGACCCGATTCGGGCTGCGCGATCTGTTT 2696
Qy 61 CTGTCGGATGAACGTCCTGCGCAAAATTAACCTCGGATTTGTTCTATAAAGATGAGACG 120
Db 2695 CTGTCGGATGAACGTCCTGCGCAAAATTAACCTCGGATTTGTTCTATAAAGATGAGACG 2636
Qy 121 GGCACAAACCCGGTACTGACAGCGTGAAGAAAGCTCAACAGTATCTGCTCGAAATGAA 180
Db 2635 GGCACAAACCCGGTACTGACAGCGTGAAGAAAGCTCAACAGTATCTGCTCGAAATGAA 2576
Qy 181 ACCACCAAAATTAACCTCGGATTTGACGGCATCCCTGAAATTTGGTCTGCTCACTCAGGAA 240
Db 2575 ACCACCAAAATTAACCTCGGATTTGACGGCATCCCTGAAATTTGGTCTGCTCACTCAGGAA 2516

```

```

Qy 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGCTCGCAGGCACAGACT 300
Db 2515 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGCTCGCAGGCACAGACT 2456
Qy 301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCCGATTTCTGCGCAAAATAACAGCGTT 360
Db 2455 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCCGATTTCTGCGCAAAATAACAGCGTT 2396
Qy 361 AAGCGTGTGGGTGAGCAACCAAGCTGCCCAACCATAGAGCGTCTTTAACTCTGCA 420
Db 2395 AAGCGTGTGGGTGAGCAACCAAGCTGCCCAACCATAGAGCGTCTTTAACTCTGCA 2336
Qy 421 GGTCTGGAAGTTCTGTTGAATAACGCTTATATGATGCGGAAATCACACTTTGACTTCGAT 480
Db 2335 GGTCTGGAAGTTCTGTTGAATAACGCTTATATGATGCGGAAATCACACTTTGACTTCGAT 2276
Qy 481 GCACTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGCTAGTGTCTGTTCCATGGCTGC 540
Db 2275 GCACTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGCTAGTGTCTGTTCCATGGCTGC 2216
Qy 541 TGCCTAAACCAACCGGTATCGACCTAGCTGGAACAATGGAACAACATGGCAAACTC 600
Db 2215 TGCCTAAACCAACCGGTATCGACCTAGCTGGAACAATGGAACAACATGGCAAACTC 2156
Qy 601 TCCGTTGAGAAAGGCTGTTTACCGCTGTTTGAATTCGCTTACCAGGGTTTGGCCGCTGGT 660
Db 2155 TCCGTTGAGAAAGGCTGTTTACCGCTGTTTGAATTCGCTTACCAGGGTTTGGCCGCTGGT 2096
Qy 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCGCGCTATGCAATGAAGAGCTGATTTGT 720
Db 2095 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCGCGCTATGCAATGAAGAGCTGATTTGT 2036
Qy 721 GCCAGTTCTTACTCTAAAGCTTTGGCTGTACAACGAGCGTGTGGCGCTTGTACTCTG 780
Db 2035 GCCAGTTCTTACTCTAAAGCTTTGGCTGTACAACGAGCGTGTGGCGCTTGTACTCTG 1976
Qy 781 GTTGCTGCCGACAGTGAACCGCTTGTATCGCGCATTCAGCCAAATGAAAGCGGATTCGC 840
Db 1975 GTTGCTGCCGACAGTGAACCGCTTGTATCGCGCATTCAGCCAAATGAAAGCGGATTCGC 1916
Qy 841 GCTAACTACTCTAACCCACAGCAGACACGCGCTTCTGTTGTCGCCACCATCTCAGCAAC 900
Db 1915 GCTAACTACTCTAACCCACAGCAGACACGCGCTTCTGTTGTCGCCACCATCTCAGCAAC 1856
Qy 901 GATCGTTAGTGGATTTGGGAACAAGAGCTGATATGCGCAGCGTATTCAGCGT 960
Db 1855 GATCGTTAGTGGATTTGGGAACAAGAGCTGATATGCGCAGCGTATTCAGCGT 1796
Qy 961 ATGCGTCAGTTGTTCTGTCATAGCTGAGGAAAGGCGCAACCGCGACTTTCAGCTTT 1020
Db 1795 ATGCGTCAGTTGTTCTGTCATAGCTGAGGAAAGGCGCAACCGCGACTTTCAGCTTT 1736
Qy 1021 ATCATCAACAGAACCGCATGTTCTCTTCAGTGGCTGACAAAGAACCAAGTCTCGCT 1080
Db 1735 ATCATCAACAGAACCGCATGTTCTCTTCAGTGGCTGACAAAGAACCAAGTCTCGCT 1676
Qy 1081 CTGCGGAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTCGCGGATG 1140
Db 1675 CTGCGGAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTCGCGGATG 1616
Qy 1141 ACACCAAGATGAACATGCTGCGCTGTGGAAGCGATTTGTCAGTGTGCTGTA 1191
Db 1615 ACACCAAGATGAACATGCTGCGCTGTGGAAGCGATTTGTCAGTGTGCTGTA 1565

```

RESULT 4

```

US-10-369-493-44812
; Sequence 44812, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.

```

```

; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 44812
; LENGTH: 1188
; TYPE: DNA
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-44812

```

Query Match	55.1%	Score 555.8	DB 6	Length 1188
Best Local Similarity	72.0%	Pred. No. 3.8e-216		
Matches 855	Conservative 0	Mismatches 332	Indels 0	Gaps 0
Qy	1	ATGTTTGAGAACATTACCGCGCTCTCTGCCAGCCGATTCTGGGCGCTGCCGATCTGTTT	60	
Db	1			
Qy	61	CGTGCCGATGAACTGCCGGGCAAAATTAACTTCGGGATTGGTGCTATAAAGATGAGACG	120	
Db	61			
Qy	121	GGCAAAACCCGGTACTGACACGCGTGAAAAGCTGAACAGTATCTGCTCGAAAATGAA	180	
Db	121			
Qy	181	ACCACAAAATTAACCTCGGCAITGACGGCATCCCTGAATTGGTCGCTGCACCTCAGAA	240	
Db	181			
Qy	241	CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCTCGCACGGCACAGACT	300	
Db	241			
Qy	301	CCGGGGGCACTGCGGCACACTACGGTGGCTGCCGATTTCCCTGGCAAAAATAACACAGCGT	360	
Db	301			
Qy	361	AAGCGTGTGGGTGAGCAACCCAAAGCTGGCCGAAACCAATAGAGCGCTTTTAACTCTGCA	420	
Db	361			
Qy	421	GGTCTGGAAGTTCTGTGAATAGCTTATTATGATGGGAAAATCACTCTTGACTTCGAT	480	
Db	421			
Qy	481	GCACTGATTAAACGCTGAATGAAGACTCAGGCTGGCAGCTAGTGTCTGTCATGGCTGC	540	
Db	481			
Qy	541	TGCCATAACCCAAACCGGTATCGACCCCTACGCTGGAAACAAATGGCAAAACACTGGCAACATC	600	
Db	541			
Qy	601	TCGGTTGAAAAGCGTGTACCGCTGTTTTCGCTTACCGGTTTTCGCGGTGTTTCCCGTGGT	660	
Db	601			
Qy	661	CTGGAAGAAGATGCTGGAAGACTCGCGCTTTTCGCGGCTATGTCATAAAGAGCTGATTGTT	720	
Db	661			
Qy	721	GCGAGTTCCTACTCTAAAACTTTGGCCCTGTACACGAGCGTGTGGCGCTTGTTACTCTG	780	
Db	721			

RESULT 5
 US-10-369-493-23918
 ; Sequence 23918, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 23918
 ; LENGTH: 966
 ; TYPE: DNA
 ; ORGANISM: *Xenorhabdus nematophilus*
 US-10-369-493-23918

Query Match	42.8%	Score 510;	DB 6;	Length 366;
Best Local Similarity	70.5%;	Pred. No. 1.6e-165;		
Matches 681;	Conservative 0;	Mismatches 285;	Indels 0;	Gaps 0
Qy	223	GGTCGCTGCAC	TACAGGAAC	TGCTTTTGGTAAAGGTAGCGCCCTGATCAATACGACAAACGT 282
Db	1	GGCCGGTGATT	CAGGAAC	TGCTTTTGGAGTACCAGTGCATCGTTACTGAAAAACGT 60
Qy	283	GCTCGACGGCA	CAGACTCCGGGGGGG	CACTGCGGCACCTACGCTGGCTGCCGATTTTCCTG 342
Db	61	GCCCGTACCGT	ACAAAGCCAGGCGG	CAACAGGCGCACTGCGTACTGCGCGCTGATTTTATT 120
Qy	343	GCAAAAAATAC	CCAGCGTTAAGCGTGTGTGGTGAGCAAC	CCCAAGCTGSCCGAACCATTAAG 402
Db	121	GCCAAAGCAACT	ATATGCCAAAC	GGTTTGGATCAGTAAACCAACTGSCCAACCATATA 180
Qy	403	AGCGTCTTTAA	CTCTCGAGGTC	TGGAAGTTCGTGAATACGCTTTATTATGATCGCGAAAAAT 462
Db	181	GGTGTTTTTTTC	CCAGCGGAGGTTT	TAGAGATCCGGAATATACTATTACGATGCGAAAAA 240

```
Qy 463 CACACTCTTGACTTCGATGCACTGAATTAACAGCCTGAATGAAGCTCAGGCTGGGACGTA 522
Db 241 CACGCTCTGAATTCGAAGGCATCTGGCAAGCCTGTCTGAAGCACAAGCGGGTGATGTT 300
Qy 523 GTGCTGTTCCATGGCTGCTGCCATAACCCCAACCGGTATCGACCTACGCTGGGAACATGG 582
Db 301 GTTCTGTGACGGTTCCTGCCATAACCCCAACCGGTATCGATCTCTCCAGAACAGTGG 360
Qy 583 CAAACACTGGCACAACCTCTCCGTTTGAGAAAGGCTGGTTACCGCTGTTTGACTTCGCTTAC 642
Db 361 CAAAACCTGGCAGATTTATCTGGGCAATAGGCTGGCTGCTGTATTCGATTCGCTTAT 420
Qy 643 CAGGTTTTGCGCGTGCTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGGCTATG 702
Db 421 CAGGGCTTTGCTAAAGTCTGGATGAAGATGCGGAAGGCTCGGCGATTTTGGCAAAAAC 480
Qy 703 CATAAAGAGCTGATTTGTCAGTTCTTACTCTTAAAACTTTGGCCTGTACAAAGCGGT 762
Db 481 CATAATGAATGATCGTTGCGAGCTCTTACTTCCAAAACCTTGGCCTGTACAAATGAGCGT 540
Qy 763 GTTGGCGCTTGACTCTGGTTGTCGCGACAGTGAACCCGTTGATCGCGATTCAGCCAA 822
Db 541 GTTGGCGCTTGACTATTTGTCGAGACAGATACCGGAGAAAGCGTTTAGTCAG 600
Qy 823 ATGAAAGCGCGATTCGCGCTAACTACTTAAACCCACGACACACGCGCTTCTGTTGTT 882
Db 601 GCTAAATTTATGTAAGTACCACTATTTCTAAACCCACGCGCTCATGGTGTCTGTTGTT 660
Qy 883 GCACCACTCTGAGCAACGATGCGTTAGCTGGGATTTGGGAAACAGAGCTGACTGATG 942
Db 661 ACACAAATTTCTGCCAAACGAAAGTCTGAAAGCAGAGTGGATTCAGGAACCTGGCAACAATG 720
Qy 943 CGCAGCGTATTCAGCGTATGCGTCACTGTTGCTCAATACGCTGCAGGAAAGGCGCA 1002
Db 721 CGTGAGCGCAATTCAGCGTATGCGCAATTTGTTGCTCAACCTTACAGAAAGAGTGCA 780
Qy 1003 AACCGGACTTCAGCTTTATCATCAACAGAACGGCATGTTCTCTTCAGTGGCCTGACA 1062
Db 781 AAACAGGATTTACGTTTCATCTGTGCTCAAAATGGTATGTTCTCATTCACGCGTCTGACT 840
Qy 1063 AAAGAACAGTGTGCTGCTGCGGAAAGATTTGGCGTATATGGGTGTTCTTGTGTCG 1122
Db 841 AAAGAACAGTGTGAACGCTGCGTGAAGAAATACGCTGTTTATGCAAGTCAAGCTCGGCGCT 900
Qy 1123 GTAAATGTGCGGAGTACACCAAGATAATGCTCCGCTGTCGGAAGCGATTGGCA 1182
Db 901 ATTAAGTCTGCTGTTTAACTGGAACATGTTCTCTGTGTGAAGCGATTGTGCT 960
Qy 1183 GTGCTG 1188
Db 961 GTACTG 966
```

```
RESULT 6
US-10-450-763-23995
; Sequence 23995, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 23995
; LENGTH: 2133
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1)-(774)
; OTHER INFORMATION: 90% homologous to Escherichia coli ATP-dependent helicase
; OTHER INFORMATION: HrpA homolog, accession number D90780, Smith-Waterman Score=1183.
US-10-450-763-23995

Query Match 41.7%; Score 496.2; DB 9; Length 2133;
Best Local Similarity 99.4%; Pred. No. 15e-160;
Matches 498; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 687 CGCTTTTCGGCGTATGCATAAAGAGCTGATGTTGTTGCCAGTTCCTACTCTAAAACTTTGG 746
Db 1425 CCCTTTTCGGCGTATGCATAAAGAGCTGATGTTGTTGCCAGTTCCTACTCTAAAACTTTGG 1484
Qy 747 CTTGTACAACGAGCGTGTGGCGCTTGTACTCTGTTGCTGCCGACAGTGAACCCGTTGA 806
Db 1485 CTTGTACAACGAGCGTGTGGCGCTTGTACTCTGTTGCTGCCGACAGTGAACCCGTTGA 1544
Qy 807 TCSCGCAATTCAGCCAAATGAAGCGCGATTTCGCGCTAACTCTTAACCCACGACGACA 866
Db 1545 TCSCGCAATTCAGCCAAATGAAGCGCGATTTCGCGCTAACTCTTAACCCACGACGACA 1604
Qy 867 CGCGCGCTTCTGTTGTTGCCACCATCTCTGAGCAACGATGCGTTACGTCGCGATTTGGGAACA 926
Db 1605 CGCGCGCTTCTGTTGTTGCCACCATCTCTGAGCAACGATGCGTTACGTCGCGATTTGGGAACA 1664
Qy 927 AGAGCTGACTGATATGCGCGAGCGTATTCAGCGTATGCGTTCAGTGTTCGTCGAATAGCT 986
Db 1665 AGAGCTGACTGATATGCGCGAGCGTATTCAGCGTATGCGTTCAGTGTTCGTCGAATAGCT 1724
Qy 987 GCAGGAAAGGCGCAACCGCGACTTCAGCTTTATCATCAACAGAACCGCATGTTCTC 1046
Db 1725 GCATGAAAGGCGCAACCGCGACTTCAGCTTTATCATCAACAGAACCGCATGTTCTC 1784
Qy 1047 CTTTCAGTGGCTGACAAAAGAAACAAGTGTGCTGCGTTCGCGAAGAGTTTGGCGTATATGC 1106
Db 1785 CTTTCAGTGGCTGACAAAAGAAACAAGTGTGCTGCGTTCGCGAAGAGTTTGGCGTATATGC 1844
Qy 1107 GGTGTTCTTCGTGCGCGTAAATGTGCGCGGATGACACGATGAACATGCGTCCGCTGTG 1166
Db 1845 GGTGTTCTTCGTGCGCGTAAATGTGCGCGGATGACACGATGAACATGCGTCCGCTGTG 1904
Qy 1167 CGAAGCGATTGTGCGAGTCT 1187
Db 1905 CGAAGCGATTGTGCGAGTCT 1925
```

```
RESULT 7
US-10-450-763-25886
; Sequence 25886, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 25886
; LENGTH: 3222
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

; NAME/KEY: SIMILAR
 ; LOCATION: (1) (774)
 ; OTHER INFORMATION: 90% homologous to Escherichia coli ATP-dependent helicase
 ; OTHER INFORMATION: HrpA homolog, accession number D90780, Smith-Waterman Score=1183.
 US-10-450-763-25886

Query Match 41.7%; Score 496.2; DB 9; Length 3222;
 Best Local Similarity 99.4%; Pred. No. 1.9e-160; Indels 0; Gaps 0;
 Matches 498; Conservative 0; Mismatches 3;

Qy	687	CGCTTTCCGGCTATGCAATGAAGAGCTGATTTGTTGCCAGTTCTCTACTCTCTAAACCTTTGG	746
Db	1425	CCCTTTCCGGCTATGCAATGAAGAGCTGATTTGTTGCCAGTTCTCTACTCTCTAAACCTTTGG	1484
Qy	747	CCTGTACACGAGCGTTGGCGCTTGTACTCTGGTTGCTGCCACAGTGAACCGTTGA	806
Db	1485	CCTGTACACGAGCGTTGGCGCTTGTACTCTGGTTGCTGCCACAGTGAACCGTTGA	1544
Qy	807	TCGGCGCATTCAGCCCAATGAAGCGCGATTCGCGCTAACTCTAAACCCACCGACACA	866
Db	1545	TCGGCGCATTCAGCCCAATGAAGCGCGATTCGCGCTAACTCTAAACCCACCGACACA	1604
Qy	867	CGGCGCTTCTGTTGTCGCCACCTCTGAGCAAGATCGGTTAGCTGCGATTTGGGAACA	926
Db	1605	CGGCGCTTCTGTTGTCGCCACCTCTGAGCAAGATCGGTTAGCTGCGATTTGGGAACA	1664
Qy	927	AGAGCTGACTGATATGCGCCAGCGTATTCAGCGTATGCGTCAATACGCT	986
Db	1665	AGAGCTGACTGATATGCGCCAGCGTATTCAGCGTATGCGTCAATACGCT	1724
Qy	987	GCAGGAAAGGCGCAACCGCGATTCAGCTTATCATCAACAGAACCGCATGTTCTC	1046
Db	1725	GCATGAAAGGCGCAACCGCGATTCAGCTTATCATCAACAGAACCGCATGTTCTC	1784
Qy	1047	CTTCAGTGGCTGACAAAGAACCAAGTCTGCGTCTGCGGAGAGTTTGGCGTATATGC	1106
Db	1785	CTTCAGTGGCTGACAAAGAACCAAGTCTGCGTCTGCGGAGAGTTTGGCGTATATGC	1844
Qy	1107	GTTTGGCTTCTGCTGCGTAAATGTGGCCGGGATGACACAGATAACTGCTCGCTGTG	1166
Db	1845	GTTTGGCTTCTGCTGCGTAAATGTGGCCGGGATGACACAGATAACTGCTCGCTGTG	1904
Qy	1167	CGAAGCGATTTGTCAGTGCT	1187
Db	1905	CGAAGCGATTTGTCAGTGCT	1925

RESULT 8
 US-10-795-159-536
 ; Sequence 536, Application US/10795159
 ; Publication No. US20050221439A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BAKALETZ et al.
 ; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
 ; FILE REFERENCE: 28335/38815A
 ; CURRENT APPLICATION NUMBER: US/10795,159
 ; CURRENT FILING DATE: 2004-03-05
 ; PRIOR APPLICATION NUMBER: US 60/453,134
 ; PRIOR FILING DATE: 2003-03-06
 ; NUMBER OF SEQ ID NOS: 771
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 536
 ; LENGTH: 7977
 ; TYPE: DNA
 ; ORGANISM: H. influenzae
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (4108)..(4108)
 ; OTHER INFORMATION: n = a, c, g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (7960)..(7960)
 ; OTHER INFORMATION: n = a, c, g, or t

US-10-795-159-536
 Query Match 41.4%; Score 493.4; DB 9; Length 7977;
 Best Local Similarity 63.4%; Pred. No. 3.1e-159;
 Matches 755; Conservative 0; Mismatches 436; Indels 0; Gaps 0;

Qy	1	ATGTTTGGAGAACATTACCGCGCTCTCGCGAGCCCGGATCTTGGCGCTTGGCGGATCTGTTT	60
Db	1594	ATGTTTGGAGAACATTAAAGCGGACCGGATCTTGGCGCTTGGCGGATCTGTTT	1653
Qy	61	CGTCCGATGAACGCTCCCGCAAAATTAACCTCGGATTTGGTCTCTATAAAGATGAGACG	120
Db	1654	AAATCCGAAACTCGCGAAATTAATTAACCTGGTATTTGGCGTTTATAAGATGCGCAA	1713
Qy	121	GGCAAAACCCGGTACTGACCGAGCTGAAAGAGCTGAACAGTATCTGCTCGAAATGAA	180
Db	1714	GGCACAAACCCGATTAATGCGTGAAGAGAGAGCAAAACGCTTATTTGATAAAGGAA	1773
Qy	181	ACCACCAAAATTTACCTCGGCATTGACGCGATCCCTGAAATTTTGGTCTGCTCACTCAGGAA	240
Db	1774	AAACCGAAATTTATCTGACTATCGATGGTATTTGCTGATTAACGAAACAAACGAG	1833
Qy	241	CTGCTCTTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTCTCGACGGCACAGACT	300
Db	1834	TTACTTTTTCGGTAAAGATTTCTGAAGTCAATAATCTAAATCGTGAAGACAGTACAAAGT	1893
Qy	301	CGGGGGGCACTGGCGCATAACGCTGGCTGCGGATTTCTTGGCAAAATATACAGGCTT	360
Db	1894	TTAGCGGAAACAGCGGCAATTAACGCTTGGCGGAGAAATTTATTAACGCTCAACTAAGCA	1953
Qy	361	AAGCGTGTCTGGGTGAGCAACCAAGCTGCGCGAACCAATAAGAGCGTCTTTAACTCTGCA	420
Db	1954	CAAAATGTTTGGATAGCACCCCACTTGGCCAAACCAATGCCATTTTCAATGGCGTC	2013
Qy	421	GGTCTCGAAAGTTCTGTAATACGCTTATATGATGCGGAAATACACACTCTTGACTTCGAT	480
Db	2014	GGTATGCAATTCGTAATATCGTTATATGATGCTGAACGCAAGCCCTTGTATGGGAT	2073
Qy	481	GCACTGATTAAACGCTGAATGAAGCTCAGGCTGGCGAGTAGTGTCTGTTCCATGGCTGC	540
Db	2074	CATTTACTTGAAGATTTAAGCCAAAGCAAGCGAGCGATGTGTTGTCTTTTACACGGTGT	2133
Qy	541	TGCCATAACCCAAACGGTATCGACCTAGCTGGACCAATGGCAACACACACACACCACTC	600
Db	2134	TGCCATAATCCGACTGCTATGACCTCTCCAGAACCAATGGCAAGAAATAGCCGCACTT	2193
Qy	601	TCGTTTGGAGAAAGGCTGGTTACCGCTGTTTGAATTCGCTTACCGAGGTTTGGCCGCTGT	660
Db	2194	TCGGCGAAACCGCTGGTTGCCACTTTTGAATTCGCTTATCAAGGTTTGGCCAAATGGA	2253
Qy	661	CTGGAAGAGATGCTGAAGGACCTGCGCGCTTTTGGCGGCTATGATTAAGAGCTGATGTTT	720
Db	2254	TTAGATGAAGATGCTTATGGCTTGGCTGCTTTTGGCGCAACCCACAAAGAAATTTAGTA	2313
Qy	721	GCGAGTCTTACTCTAAACCTTTGGCTGTACACAGCGGTGTTGGCGCTTCTACTCTG	780
Db	2314	GCGAGTCTTCTCGAAACCTTTGGTATTAATGAACGCTGTTGGTCTTCTTACTCTT	2373
Qy	781	GTTGCTGCGCACAGTCAAAACCGTTGATTCGCGCATTCAGCCAAATGAAGCGCGATTCGC	840
Db	2374	GTGGCGAAATTCAGAAATTTGATCAACCGCATTAACCCAGTGAATCAATATTCGT	2433
Qy	841	GCTAACTACTCTAACCCACGACACAGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC	900
Db	2434	ACACTCTACTCTAAACCTGATCTCACGGCGGCAACTGTAGCAACAGTATTAATGAC	2493
Qy	901	GATGCGTTAGTGGGATTTGGGAAACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT	960
Db	2494	ACTCAACTTCGCAAGAGTGGGAAATGAATTAACCTGAAATGCGCGAAGCTATCAAAAA	2553
Qy	961	ATGCGTCAAGTTGTTGCTCAATACGCTGAGGAAAAAGGCGCAAAACCGCGACTTTCAGCTT	1020
Db	2554	ATGCGTCAATTTATTCGTTTAAAGAAATACGTTGCGGAAACAGATTTTCAGCTT	2613


```

; FILE REFERENCE: GJE-6436
; CURRENT APPLICATION NUMBER: US/10/275,026A
; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: PCT/GB01/02003
; PRIOR FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: GB 0011108.8
; PRIOR FILING DATE: 2000-05-08
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 187
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1191)
US-10-275-026A-187

Query Match          39.6%; Score 471.8; DB 7; Length 1191;
Best Local Similarity 62.3%; Pred. No. 3.3e-152;
Matches 740; Conservative 0; Mismatches 447; Indels 0; Gaps 0;

QY 2 TGTTTGAGAACATTACCGCGCTCTCCGACCGGATTTGGGCGCTGGCGGATCTGTTTC 61
DB 5 TCTTCAAGCACATCGAAGCCGCCCGCGCATCGGATCTCGTTTGGGCGAAGCATTC 64
QY 62 GTCCGAGTAACGTCGCGGCAAAATTAACCTCGGAGTTGTTGTTATAAAGATGAGACGG 121
DB 65 AAGCGGAAACCCGCCCGGCAAAAGTCAACCTCGGCATCGGCGTGCATCAAGAGCGCATCCG 124
QY 122 GCAAAACCCGGTACTACACGAGGTGAAAGGCTGNACAGTATCTGCTCGAAATGAAA 181
DB 125 GCGCGACACCTATGTGTAAGCGGTCAAGAGCCGCAAGAAACCGCTGTTGGAAAGCGAAA 184
QY 182 CCACCAAAATTTACCTCGGATTCAGCGCATCCTGAAATTTGGTGGCTGCACTCAGGAAC 241
DB 185 CAATAAAATCTACCTCACCATCGACGCGTGGCGACTACACGAGCAACCCAAATTC 244
QY 242 TGCTGTTTGTGAAGTAGCGGCTGATCAATGACAAACGTTGTCGCAOGGCAAGACTC 301
DB 245 TGCTGTTTGGCAAGACCAAGCAATCATCGCAGCGCTCGCGCAAAACAGCGCAAGCC 304
QY 302 CGGGGGCACTGGGCGACTAGCGGTGCTGCGGATTTCTGGCAAAATTAACAGGTTA 361
DB 305 TTGCGGTAACAGGCGCATTCGTTATTCGCGCGAGTTTGCCAAACGTCATGTAACGCGC 364
QY 362 AGCGTGTTGGGTGAGCAACCAAGCTGGCGCAACCAAGAGCGTCTTTAACTCTGCGAG 421
DB 365 AAACCATCTGGATTTCCAAATCCGACTTGGCCCAACCAAGCGCATCGCCAAAGCGTGC 424
QY 422 GTCTGGAAGTTTGTGAAATAGCTTTATATGATGCGGAAATACACTCTTTGACTTCGATG 481
DB 425 GTATCCAAAGACCAACCTTATCGCTACTATGATGCGCGCAAAACAGCGTTTGGATTGGGACG 484
QY 482 CACTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGAGCTAGTCTGTTCCATGGCTGCT 541
DB 485 GTATGATTTGAAGACTTGAAGCAAGCGCAAAAGGCGCATCGTCTCTGCTGACGCGTGT 544
QY 542 GCCATACCCAAACCGGTATGACCTACGCTAGCTGGAACAATGCAAACTGCGCAACTCT 601
DB 545 GCCCAACCCCTACCGGTATGACCTACGCGCAACCAATGGAAACTCTGGCAAACTTT 604
QY 602 CCGTTGAGAAAGGTGTTACCGCTGTTTGAATTCGCTTACCGGGTTTTGCCCGTGGTC 661
DB 605 CTGCGGAAAGGCTGTTACCGCTGTTTGAATTCGCTTACCGAGGCTTCGGCAATGTT 664
QY 662 TGAAGAAGATGCTGAAGACTGCGCTTTTCGGCGCTATGCATATAAGAGCTGATTTG 721
DB 665 TGAAGAAGATGCTGACGCGCTGCGCGTGTCTTGAACAACAAATACAGAAATTTGCTGATTG 724
QY 722 CCAGTTCTCTTAAAACTTTGGCTGTACACGAGCGTGTGGCGCTGTGACTCTGG 781
DB 725 CCAGCTCTTATCCAAAACCTTCGTTATGTACACGAGCGGTGCGCGCTTCACTTTGG 784
```

```

QY 782 TTGCTGCCAGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGCGGATTCGCG 841
DB 785 TGGCCGAAGATGAAGCAACCGCGCCGCCCAAGTCAAAACCATTTATCCGTA 844
QY 842 CTAACCTACTCTAAACCCACAGACGCGGCTTCTGTTGTTGTCACCATCTCTGAGCAACG 901
DB 845 CCTTGTTATTTCCAAACCGCTTTCACACGCGGCGCAACACTATTGGCTGGTGTGAAAAATG 904
QY 902 ATGCGTTACGTGGGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTCAGCGTA 961
DB 905 ATGATTTGAAAGCACAATGGATTGCCGAACTCGATGAAATGCGCGCGCGTATCAAGGCCA 964
QY 962 TCGCTCAGTTGTTGTCATACGCTGCAGGAAAAAGCGCAACCGCGACTTCAGCTTTA 1021
DB 965 TGGCCCAAAATTTGTCGAGTTGCTCAAGCCAAAGTGCNACTCAAGACTTTGATTTCA 1024
QY 1022 TCATCAAAACAGAACGCGCATGTTCTCTTCAGTGGCTGCAAAAGAACAAAGTCTCGCTC 1081
DB 1025 TTATCGAACAAAACGCGCATGTTCTCTTCAGCGGCTTGACTCCCGAAACAAGTCGACGCTT 1084
QY 1082 TGGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGCTAAATGTTGGCCGGATGA 1141
DB 1085 TAAAAAACGAGTTTGCAATTTATGCGGTCCTCGCGCGCATCAACGTCGCGCGCATTA 1144
QY 1142 CACCAGATAACATGGCTCGCTGTGCAAGCGATTGTGCGAGTGCTG 1188
DB 1145 CGAAGCAACATCGATTATCTGTGCAAGTATCGTAAAGTACTG 1191
```

```

RESULT 11
US-10-915-740A-40
; Sequence 40, Application US/10915740A
; Publication No. US20050191316A1
; GENERAL INFORMATION:
; APPLICANT: Frazer, Claire M.
; APPLICANT: Hickey, Erin
; APPLICANT: Peterson, Jeremy
; APPLICANT: Tettelin, Herve
; APPLICANT: Venter, J. Craig
; APPLICANT: Maignani, Vega
; APPLICANT: Galeotti, Cesira
; APPLICANT: Mora, Manroza
; APPLICANT: Ratti, Giulio
; APPLICANT: Scarselli, Maria
; APPLICANT: Scarlato, Vincenzo
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
; FILE REFERENCE: 002441.00090
; CURRENT APPLICATION NUMBER: US/10/915,740A
; CURRENT FILING DATE: 2004-08-11
; PRIOR APPLICATION NUMBER: 09/806,866
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: USSN 60/103,794
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: USSN 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/US99/25373
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 1068
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 17381
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-915-740A-40
```

Query Match 39.2%; Score 466.8; DB 9; Length 17381;
Best Local Similarity 62.0%; Pred. No. 8.3e-150;
Matches 738; Conservative 0; Mismatches 452; Indels 0; Gaps 0;


```
Db 562674 TCGCGCGTACGGCGCATTCGGTATTTCGGCCGAGTTTGCCAAACGCCAGTTGAACGCGC 562615
Qy 362 AGCGTGTGGGTGAGCAACCCAGCTGGCGGAAACCAATAGAGCGGTCTTTAACTCTGCAG 421
Db 562614 AAACCATCTGGATTTCGAATCCGACTTGGCCCAACCAACGCGCATCGCCAAAGCGGTGC 562555
Qy 422 GTCTGGAAGTTCGTGAATACGCTTATTATGATCGGGAATATCACACTCTTGACTTCGATG 481
Db 562554 GTATCCAAAGCAAAACCTTATTCGTTACTATGATCCGCAACACGCTTGGATGGGACG 562495
Qy 482 CACTGATTAAACAGCTCGTAATGAAGCTCAGGCTGGCGACGTAAGTCTGCTTCATATGGCTGCT 541
Db 562494 GCATGATTGAGGACTTGGCCAGCGGCAAAAGGCGCATCGTCTGCTGCACGGTGTCT 562435
Qy 542 GCCATTAACCAACCGGTATCGACCTTACGCTGCGTGAACCAATGGCAACACTGGCAACACTCT 601
Db 562434 GCCACAATCTACCGGCATCGACCTTACGCTGCGGCGCAACATGGGAACCTTTGGCAAACTTT 562375
Qy 602 CCGTTGGAAGAGCTGGTTACCGCTGTTTGAATCTCGCTTACAGGGTCTTGGCCCGGTGTC 661
Db 562374 CTGCCGAAAAGGCTGGTTCCCGCTGTTCGCTTTCGCTTACCAAGGCTTCGGCAATGGTT 562315
Qy 662 TCGAAGAGATGCTGAAGAGCTGCGGCTTTCGCGGCTATGCATAAAGAGCTGATTGTTG 721
Db 562314 TGGAGAGATGCTACGGCTTGGCGGTCTTCTTGAAACACATACAGATTCTGATG 562255
Qy 722 CCAGTTCCTACTCTAAACCTTTGGCGCTGTACACAGCGGTGTGGCGCTTGTACTCTG 781
Db 562254 CCAGCTCTTATTCAAAACCTTCGGTGTGTACACAGCGGTTCGCGCGCTTCACTTTGG 562195
Qy 782 TTGCTGCGCAGTGAACCGTGTGATCGCGCATTCAGCCAAATGAAGCGGCGATTCGCG 841
Db 562194 TGGCCGAAGATGAAGAAACAGACGCGCGCCCAACAGCAAGTCAAAACCATCATCCGTA 562135
Qy 842 CTAACACTCTAACCCACAGCACAGCGGCTTCTGTGTTGCGCACCATCTCAGCAACG 901
Db 562134 CTTTGTATTCCAACCCGGCTTCAACGCTGCGAACCATTCGCGTGTGGTGTGAAATG 562075
Qy 902 ATGCGTTACGTGGGATTGTGGAACAAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGTA 961
Db 562074 ATGATTGAAGACAACTGATTGCCGAATCGATGAATGCGGCGCGCATCAAGCCA 562015
Qy 962 TGGCTGATGTTGCTCAATACGCTGAGGAAAGGCGGCAAAACCGGACTTCAGCTTTA 1021
Db 562014 TGGCGCAAAATTTGTCGGGTGCTCAAGCCAAAGGTGCAAGCCAAACTTTGATTTC 561955
Qy 1022 TCATCAACAGACGCGATGTTCTTCACTAGTGCCTTGACAAAGAACAGTCTGCGTC 1081
Db 561954 TTATCAACAAACGGTATGTTCTTTTACGGCTTGAATCCCGAACAAAGTCGACCGCC 561895
Qy 1082 TGGCGGAAGATTGGCGTATATGCGGTTGCTTCTGTCGCTAAATGTGGCCGGATGA 1141
Db 561894 TGAACACGAGTTTGCAATTTATGCGCTCGCTCCGCGCATCAAGTCCCGGCATTA 561835
Qy 1142 CACAGATAACATGGCTCCGCTGTGGAACGCAATTTGGCAGTGTGTA 1191
Db 561834 CCGACAACAACATCATCTATCTGCGAAAGCATCGTGAAGTACTGTAA 561785

RESULT 13
US-10-329-670-1
; Sequence 1, Application US/10329670
; Publication No. US20040018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; FILE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
```

```
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
```



```
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152530)..(152530)

Query Match      38.5%; Score 458.2; DB 7; Length 1830121;
Best Local Similarity 61.5%; Pred. No. 1.1e-145;
Matches 733; Conservative 0; Mismatches 458; Indels 0; Gaps 0;

Qy      1  ATGTTTGAGAACATTACCGCGCTCTGCGACCCCGATTCTGGCGCTGGCGGATCTGTTT 60
Db      1684147  ATGTTTGACATATCAAAGCGGCACCGCATCTTAGCTTAGCGAAGCATTT 1684206

Qy      61  CGTGCCGATGAACGTCCCGCAAAATTAACTCGGGATTGGTGTCTATAAAGATGACACG 120
Db      1684207  AAATCCGAAACTCGCGAAATAAAATCAATTTGGGTATTGGCGTTTATAAAGATGCCAA 1684266

Qy      121  GGCACAAACCCGGTACTGACCGACGCTGAAAGCTCAACAGTATCTGCTCGAAATGAA 180
Db      1684267  GGCACAAACCCCAATTATGACGCGGTAAAGAGCCGAAACAGTATTTTGATAGGAA 1684326

Qy      181  ACCACCAAAATTTACCTCGGCATTGACGGCATCCCTGGAATTTGGTCCCTGCACCTCAGGAA 240
Db      1684327  AAAACCAAGAAATTATCTGACTATCGATGTTTGGCGATTATAACGAACAAACAAAGCA 1684386

Qy      241  CTGCTGTTTGGTAAAGGTAGCGCCTGATCAATGACAAAGTGTCTCCAGCGCACACT 300
Db      1684387  CTCCTTTTCGTTAAAGATTCTGAAGTTCATCCAATCTAATCGAGCAAGACAGTACAAAGT 1684446

Qy      301  CCGGGGGGCACGCGCACACTACGCTGCGGCTGCGGATTTCCTGGCAAAATATACCAGCGTT 360
Db      1684447  TTAGGCGGAACAGGTGCAATTACGCATTGCGGAGAAATTTTAAACCCCAAACTAAAGCA 1684506

Qy      361  AAGCGTGTGTGGGTGAGCAACCCCAAGCTGGCCCAACCATAGAGCGTCTTTAACTCTGCA 420
Db      1684506  AAGCGTGTGTGGGTGAGCAACCCCAAGCTGGCCCAACCATAGAGCGTCTTTAACTCTGCA 420
```

```

Db 1684507 CAAAAATGTTGGWTCAGCACACCAAACTTGCCAAAACCAATGCGATTTTCAATGCTGTC 1684566
Qy 421 GGTCTGGAAGTTCGTGATACGCTTATTATGATGCGGAAATACACTCTTGACTTCGAT 480
Db 1684567 GGTATGACCAATTCGTGAATATCGTATTATATGATGCTGAAGCGCAAGCCCTTGATGGGAA 1684626
Qy 481 GCACCTGATTAAACAGCCCTGAATGAAGCTFCAGCTGCGGACCTAGTGCCTGCTTCCATGGCTGC 540
Db 1684627 CACTTATTAGAGATTTTAAGCCAAGCAAGCAAGGCGATGTGTGCTTTTACACGGTTGT 1684686
Qy 541 TGCATTAACCCAAACCGGTATGACCCCTACGCTGGAACAATGGAACAACACTGGCACAACTC 600
Db 1684687 TGCCATTAATCCGACTGCTGATTGACCCCTACTCCAGAACATATGCGCAAGAAATAGCGCACTT 1684746
Qy 601 TCGGTTAGAAAGCTGGTTACCGCTGTTTACGCTTTCGCTTACCAGGCTTTTGCCCGTGGT 660
Db 1684747 TCAGCTAAAAATGGTTCGGTTGCCACTCTTTGCACTTTTCAAGGTTTATAGCAACGGA 1684806
Qy 661 CTGGAAGAAGATGCTGAAGGACTGCGGCTTTTGGCGCTATGCTAAGAGCTGATTGTT 720
Db 1684807 TTATGATGAAGATGCTTATGTTTACGTGCTTTTGACGAAACCAACCAAGAAATATTAGTG 1684866
Qy 721 GCCAGTTCTCTACTATAAACTTTGGCCTGTACAAAGAGCGTGTGGCGCTTGTACTCTG 780
Db 1684867 GCGAGTTCACTTCGAAAACTTTGGTTTATATAATGAACGTCGTGTGTGCACTTACCCCTT 1684926
Qy 781 GTTGCTGCCAGAGTGAACCGTTGATCGGCGATTTAGCCAAATGAAGCGCGATTCGC 840
Db 1684927 GTGCGAAAAATGCAGAAATTTGCATCAACCTCATTAACCAAGTAAATCAATTTATTCGC 1684986
Qy 841 GCTAACTACTCTAACCCACAGACACGCGCTTCTGTTGTCACCATCTCTGAGCAAC 900
Db 1684987 ACCCTATACTCTAACCCAGCTTCTCAGCGCGGGCGACCGTAGCAACAGTATTAAATGAT 1685046
Qy 901 GATCGGTTAGCTGCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCGT 960
Db 1685047 GCTCAACTTCGCCAAGATGCGAATGAATTAACAGTAAATGCGTGACGCATCAAAAA 1685106
Qy 961 ATGCGTCAGTTGTTGTCATACGCTGCAGGAAAAAGCGCAACCGCGACTTCAGCTTT 1020
Db 1685107 ATGCGTCACATTAATTCGTTCACTGTTATTAAGAAATATGTCAGAAACAAGATTTTCA 1685166
Qy 1021 ATCATCAACAGACGCGCATGTTCTCCTTCAGTGGCTGACAAAGAACAAAGTGCCTG 1080
Db 1685167 ATCATCAACAAACGCGTATGTTCTTCTCAGTGGATTAACAGGGAAACAGTGGATCGT 1685226
Qy 1081 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGCGTAAATGTGGCGCGGATG 1140
Db 1685227 TTAAGAAATGAATTTGCCATTTACGCTGTTCTGTTCTGTCATCAACGTAGCTGGAATC 1685286
Qy 1141 ACACGATTAACATGGCTCCGCTGTGCGAAGCGAATTTGCGAGTGCCTGTAA 1191
Db 1685287 ACAGAGATAATATTCGCTATCTATGTGAAAGTATCGTGAAGTACTTTAA 1685337

RESULT 14
US-10-158-865-1
; Sequence 1, Application US/10158865
; Publication No. US20040203093A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCES: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1

```

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g

```



```
Db 1684567 GGTATGACCATTCGTGAATATCGTTATTATGATGCTGAACGCAAGCCCTTGATGGGAA 1684626
Qy 481 GCACGTGATTAAAGCCGTAATGAAGCTCAGCGCTGGCGACGCTAGCTGTTCCATGCTGC 540
Db 1684627 CACTTATTAAGAATTTAAAGCAAGCAAGCGGATGTGTGCTTTTACACGGTGT 1684686
Qy 541 TGCATAAACCAACCGGTATCGACCCCTACGCTCGAACCAATGGCAAAACACTGGGCACAACCTC 600
Db 1684687 TGCATTAATCCGACTGGTATTGACCCCTACTTCAGAACATGGCAAGAAATTAGCGCACCTT 1684746
Qy 601 TCGGTCAGAAAGCGCTGTTACCGCTGTTTGACTTCGCTTTACCAGGGTGTTCGCCGCTGGT 660
Db 1684747 TCAGTAAATAATGGTTGGTTGCCACTCTTTGACTTTATCAAGGTTTAGCCAAACGGA 1684806
Qy 661 CTGGAAGAAATGCTGAAGCACTCGCGCTTTTCGGGCTATGATTAAGAGCTGATTGTT 720
Db 1684807 TTAGATGAAGATGCTTATGCTTTTACGCTGCTTTTGCAGCAAAACCAAGAAATATTAGTG 1684866
Qy 721 GCCAGTTCCTACTCTAAATACTTTGGCTCTGACAAAGAGCGTGTGGCGCTTGCTACTCTG 780
Db 1684867 GCGAGTTCATCTCGAAATACTTTGGTTTATTAATGAACGTGTGTGTCATTTACCCCTT 1684926
Qy 781 GTTGCTGCCACAGTGAACCGTTTGATCGCGCATTCAGCCAAATGAAGCGGCGATTCGC 840
Db 1684927 GTGCAGAAATGCAGAAATTCATCAACCTCATTAACCAAGTAAATCAATTATTCCG 1684986
Qy 841 GCTAACTACTCTAACCCACGACACAGCGGCTTCTGTTGTCACCATCTCCTGAGCAAC 900
Db 1684987 ACCCTATACTCTAACCCAGCTTCTCACGGGGGGCGACCGTAGCAACAGTATTAAATGAT 1685046
Qy 901 GATCGGTAGTGGATTTGGACACAGAGCTGACTGATATGCGCAGCGTATTACGCGT 960
Db 1685047 GCTCAACTTCGCCAAGAAATGGAAATGAATTAACCTGAATGCGTGAACGCATCAAAAA 1685106
Qy 961 ATGCGTCAGTGTGTCGTCAATACGCTGCAGGAAAAAGCGGCAACCCGCACTTCAGCTTT 1020
Db 1685107 ATGGTCACATTATTCGTTTCAGTTATTAAGAATATGGTCAGAACAAAGATTTCAGCTTT 1685166
Qy 1021 ATCATCAACAGAACCGCATGTTCTCTTTCAGTGGCTGACAAAGAACAAAGTGTGCGT 1080
Db 1685167 ATCATTAACAAACCGTATGTTTCTTTCAGTGGATTAAACAGGGGAACAAAGTGGATCGT 1685226
Qy 1081 CTGCGCAAGATTGGCGTATATGCGGTGCTTCTGTCGCTGAATGTGGCGGGATG 1140
Db 1685227 TTAATAAATGAATTTGCCATTTACGCTGTTTCTGCTGCTATCAACGCTAGCTGGAATC 1685286
Qy 1141 ACACCATATAACATGGCTCGCTGTGCGAAGCGATTGTGGCAGTGTCTGTA 1191
Db 1685287 ACAGAAGATATATTCGCTATCTATGTGAAGTATCGTGAAGTACTTTTAA 1685337

RESULT 15
US-10-981-687-1
; Sequence 1, Application US/10981687
; Publication No. US20050131222A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2CID12
; CURRENT APPLICATION NUMBER: US/10/981,687
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US 10/158,865
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
```

```
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
```

```

; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (139910)..(139910)

```

```

; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:

Query Match      38.5%; Score 458.2; DB 9; Length 1830121;
Best Local Similarity 61.5%; Pred. No. 1.1e-145;
Matches 733; Conservative 0; Mismatches 458; Indels 0; Gaps 0;

Qy 1 ATGTTTGAGAACATTACCGCGCTCTGCGGACCCGATTCGGGCGCTGGCGCGATCTGTTT 60
Db 1684147 ATGTTTGAGAACATTACCGCGCTCTGCGGACCCGATTCGGGCGCTGGCGCGATCTGTTT 1684206
Qy 61 CGTGCCGATGAACGTCCCGGCAAAATTAACCTCGGGATTGGTGTCTATAAGATGAGACG 120
Db 1684207 AATCCGAACTCGGAAATAAATCAATTTGGGTATTGCGTTTATAAGATGCGCA 1684266
Qy 121 GGCAAAACCCCGGTACTGACGCGTGAAAGGCTGAACAGTATCTGCTCGAAATGAA 180
Db 1684267 GGCACACCCCAATTATGACGCGGTAAAGAGCCGAAACAGTATTATTGATAAGGAA 1684326
Qy 181 ACCACCAAAATTTACCTCGGGAATTGACGGGATCCCTGAAATTTGGTGTGCTGCACTCAGAA 240
Db 1684327 AAAACCAAGAAATTTATCTGACTATCGATGTTGCGGATTATACGAAACAAACAAAGCA 1684386
Qy 241 CTGCTGTTTGGTAAAGTAGCGGCTGATCAATGACAAACGTCCTCCACGGCAGACT 300
Db 1684387 CTCCTTTTCGTTAAAGATTCCTGAATCTCAATCTTAATCGAGCAAGACAGTACAAAGT 1684446
Qy 301 CCGGGGGGCACTGGCGCACTACGCTGGCTGCGGATTTCTCTGGCAAAATAATACCAGCGTT 360
Db 1684447 TTAGCGGAACAGGTGCATTACGCGGAGAAATTTATTAACGCCAACTAAAGCA 1684506
Qy 361 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGGAAACCATAGAGCGCTCTTTAACTCTGCA 420
Db 1684507 CAAAATGTTTGGWTCAGCACACCAACTTGGCCAAACCAACATGCGATTTTCAATGCTGTC 1684566
Qy 421 GGTCTGGAAGTTTGTGATAGGCTTATTATGATGCGGAAATACACTCTTGACTTCGAT 480
Db 1684567 GGTATGACCAATTCGTGAATATCGTTATTATGATGCTGAACCAAGCCCTTGATGGGAA 1684626
Qy 481 GCACTGATTAAACAGCCCTCGAATGAAGCTCAGCTGGCGACGTAGTGTGTTTCCATGGCTGC 540

```

Db 1684627 CACTTATTAGAGATTTAAGCCAAAGCAAGCGGAGTGTGGTGTCTTTTACACGGTGTG 1684686
Qy 541 TGCCTAAACCCACCGGTATCGACCCCTACGCTGGAAACATGGCAACACACTGGGCACAACTC 600
Db 1684687 TGCCTAAATCCGACTGGTATTTGACCCCTACTCCAGAACAAATGGCAAGAAATTAGCCGCACTT 1684746
Qy 601 TCCGTTGAGAAAGGCTGGTTTACCGCTGTTTGACTTTCGCTTACCAGGGTTTTGCCCGTGGT 660
Db 1684747 TCAGCTAAATAATGGTTGGTTGCCACTCTTTGACTTTTGTCTTATCAAGGTTTAGCCAACGGA 1684806
Qy 661 CTGGAAGAAGATGCTGAAGGACTCGCGCTTTTCGCGCTATGCAATAAGAGCTGATGTGT 720
Db 1684807 TTAGATGAAGATGCTTATGCTTTTACGTGCTTTTGCAGCAAAACCAAGAAGAAATATTAGTG 1684866
Qy 721 GCCAGTTCCTACTCTAAATACTTTGGCCTGTACAAAGAGCGTGTGGCGCTTGTACTCTG 780
Db 1684867 GCGAGTTTCATTCTCGAAAAAATTTGGTTTATATAATGAACGTTGGTGCATTTACCCCTT 1684926
Qy 781 GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGGCGATTTCGC 840
Db 1684927 GTGCAGAAATTCGAGAAATTTGATCNACTCATTAACCCAGTAAATCAATTAATTCGC 1684986
Qy 841 GCTAACTACTCTAACCCACAGCACACGCGCTTCTGTGTTGTCACCAATCCTCGAGCAAC 900
Db 1684987 ACCCTATACTCTAACCCAGCTTCTCAGCGCGGGCGACCGTAGCAACAGTATTAAATGAT 1685046
Qy 901 GATGCGTTACGTGGGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960
Db 1685047 GCTCAACTTCGCCAAGAATGGGAAAATGAATTAACCTGAATGGGTGAACGCATCAAAAA 1685106
Qy 961 ATGCGTCAGTTGTTTCGTCAATACGCTGCAGGAAAAAGGCGCAAAACCGCACTTCAGCTTT 1020
Db 1685107 ATGGTCACTTATTCGTTTCAGTTATTAAGAATATGGTGAGAAACAAGATTTTACGCTTT 1685166
Qy 1021 ATCATCAAAACAGAACGGCATGTTCTCCTTCAGTGGCCCTGACAAAAGAACAAAGTGTGCGCT 1080
Db 1685167 ATCATTGAACAAACCGTATGTTTCTTTCAGTGGATTAAACAGGGGAAACAAGTGGATCGT 1685226
Qy 1081 CTGGCGGAAGAGTTTGGCGTATATGCGGTTGCTTCTGGTCCGTAATATGTCGGCGGATG 1140
Db 1685227 TTAATAAATGAATTTGCAATTTACGCTGTTCTGTTCTGTCGTATCAACGCTAGCTGGAATC 1685286
Qy 1141 ACACCAAGATAACATGGCTCGCTGTGGAAGCGATTGTGGCAGTGTGTAA 1191
Db 1685287 ACAGAGATAATATTCCCTATCTATGTGAAAGTATCGTGAAGTACTTTAA 1685337

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 03:03:01 ; Search time 1180 Seconds
(without alignments)
2236.626 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045
Sequence: 1 MFENITAPADPILGLDLF.....VAGTPDNMAPLCEAIVAVL 396

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n model -DEV=xlp

-Q=/abss/ABSSWEB_spool/US10673786/runat_14032006_135608_22579/app_query.fasta_1
-DB=N Geneseq -QFMT=fastap -SUFFIX=eng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02p
-USER=US10673786 @CGN 1 1 1929 @runat_14032006_135608_22579 -NCPUS=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DEBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 21.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2045	100.0	1191	9	ADB83284 Escherich
2	2045	100.0	1191	13	AdS46042 Bacterial
3	2045	100.0	1191	14	AdW95403 Nucleotid
4	2045	100.0	1331	2	AAV40259 Escherich

5	2045	100.0	3659	1	AAN71109
6	2045	100.0	14759	4	AAS46273
c	7	2039	99.7	1293	1 AAN71108
8	1931	94.4	1209	11	ACH99107
9	1682	82.2	1191	10	ACF68649
10	1682	82.2	110000	10	ACF67367_14
11	1682	82.2	249878	10	ACF65381
12	1654	80.9	1188	13	ADT46374
13	1343	65.7	7977	13	ADT05500
c	14	1343	65.7	349980	13 ADT05649
15	1325	64.8	110000	2	AAT42063_16
16	1318	64.4	966	13	AD45488
17	1317	64.4	1269	10	ABZ40973
18	1311	64.1	1191	6	AAS97289
19	1310	64.1	17381	3	AA81493
c	20	1310	64.1	110000	3 AA81490_05
c	21	1310	64.1	349980	3 AAF21607_
22	1005.5	49.2	1185	13	ADS59837
23	996	48.7	1188	13	ADS55510
24	996	48.7	1194	13	ADS49995
25	982	48.0	1191	13	ADS56437
26	982	48.0	1242	11	ABD13824
c	27	982	48.0	1329	11 ABD13558
c	28	982	48.0	24417	2 AAT97221
29	975	47.7	1191	13	ADS63841
30	975	47.7	1191	13	ADS64221
31	975	47.7	1194	13	ADS63473
c	32	955	46.7	579	11 ACH99159
33	947	46.3	1194	13	ADS62038
34	924	45.2	1185	13	ADS57471
35	917	44.8	1185	13	ADS57235
36	917	44.8	1481	13	ADX64272
37	914	44.7	1489	13	ADX45580
38	913	44.6	1456	13	ADX47271
39	911.5	44.6	1185	13	ADS61736
c	40	911	44.5	1508	13 ADB60438
41	909.5	44.5	1509	11	ABD14364
42	909.5	44.5	1632	11	ABD13903
43	909	44.4	1203	13	ADT42736
44	908	44.4	702	10	ADF03520
45	893	43.7	1188	13	ADS56672

ALIGNMENTS

RESULT 1
ADB83284
ID ADB83284 standard; DNA; 1191 BP.
XX
AC ADB83284;
XX
DT 04-DEC-2003 (first entry)
XX
DE Escherichia coli aspartate aminotransferase gene.
XX
KW ds; gene; aspartate aminotransferase; threonine; fermentation.
XX
OS Escherichia coli.
XX
PH Key Location/Qualifiers
FT CDS 1..1191
FT /tag= a
FT /product= "aspartate aminotransferase"
XX
PN WO2003072786-A1.
XX
PD 04-SEP-2003.
XX
PF 25-FEB-2003; 2003WO-JP002067.
XX
PR 27-FEB-2002; 2002RU-00104983.
XX
PA (AJIN) AJINOMOTO CO INC.

XX Akhverdian VZ, Savrasova EA, Kaplan AM, Lobanov AO, Kozlov YI;
 XX WPI; 2003-721777/68.
 XX P-PSDB; ADB83285.
 XX Industrial production of L-threonine by fermentation using *Escherichia*
 PT modified to enhance aspartate aminotransferase activity, with improved
 PT productivity.
 XX Claim 7; Page 18-20; 26pp; Japanese.
 XX The invention relates to a bacterium belonging to the genus *Escherichia*
 CC which is modified to enhance aspartate aminotransferase activity for the
 CC production of L-threonine. The bacterium is used for the industrial
 CC production of L-threonine by fermentation. This sequence corresponds to
 CC the *E. coli* aspartate aminotransferase gene.
 XX
 SQ Sequence 1191 BP; 291 A; 305 C; 317 G; 278 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 8,796-203 Length: 1191
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 9 Gaps: 0
 US-10-673-786A-2 (1-396) x ADB83284 (1-1191)
 QY 1 MetPheGluAenIleThrAlaAlaProAlaAsePheLeuGluLeuAlaAsePhe 20
 DB 1 ATGTTTGAGAACATTACCGCGCTCTCGCCAGCCGATCTGGGCTGGCGATCTGTTT 60
 QY 21 ArgAlaAsePheGluAsePheGlyLeuAsePheGlyLeuGluAsePheGluThr 40
 DB 61 CGTCCGATGAACGTCGCGGAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 120
 QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAseGlu 60
 DB 121 GGCAAAACCCGGTACTACGACGCTGAAAGGCTGAACAGATCTGCTCGAAATGAA 180
 QY 61 ThrThrLysAenTyrLeuGlyLeuAsePheGlyLeuProGluPheGlyArgCysThrGlnGlu 80
 DB 181 ACCACAAAATTTACCTCGCATTCGCGCATCCCTGAATTTGGTGTGCTGACCTCAGGAA 240
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuLeuAsePheLysArgAlaAseThrAlaGlnThr 100
 DB 241 CTGCTGTTTGGTAAGGTAGCGCCCTGATCAATGACAAACGTGCTCGCACGGCACAGACT 300
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAsePheLeuAlaLysAseThrSerVal 120
 DB 301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCCGATTTCCCTGGCAAAAATACCAAGCTT 360
 QY 121 LysArgValTrpValSerAseProSerTrpProAseHisLysSerValPheAseSerAla 140
 DB 361 AAGCGTGTGGTGAGCAACCCAGCTGGCGCAACCAATAGAGCGCTTTTAACCTGCA 420
 QY 141 GlyLeuGluValArgGluTyrAlaTyrTrpAsePheAlaGluAseHisThrLeuAsePheAse 160
 DB 421 GGTCTGGAAGTTCTGGAATACGCTTATTTATGATGCGGAAATTCACACTCTTGACTTCGAT 480
 QY 161 AlaLeuLeuAseSerLeuAseGluAlaGlnAlaGlyAsePheValLeuPheHisGlyCys 180
 DB 481 GCACTGATTAACAGCTGTAATGAAGCTCAGCTGGCGACGATGCTGTTCCATGGCTGC 540
 QY 181 CysHisAseProThrGlyLeuAseProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
 DB 541 TGCCATAACCCAAACCGGTATCGACCTACGCTGGAAACAAATGGCAAAACACTGGCAACCTC 600
 QY 201 SerValGluLysGlyTrpLeuProLeuPheAsePheAlaTyrGlnGlyPheAlaAseGly 220
 DB 601 TCCGTTTGAGAAAGGCTGTTTACCGCTGTTTGACTTCGCTTACCAAGGTTTTTCCCGCTGTT 660

QY 221 LeuGluGluAsePheAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 DB 661 CTGGAAGAAGATGCTGAAGACTCGCGCTTTCGCGCTATGCATAAAGAGCTGATTGTT 720
 QY 241 AlaSerSerTyrSerLysAsePheGlyLeuTyrAseGluArgValGlyAlaCysThrLeu 260
 DB 721 GCCAGTTTCTTACTCTAAAACTTTGGCTGTGTACAAAGAGCGGTGTGGCGCTTGACTCTG 780
 QY 261 ValAlaAlaAsePheSerGluThrValAsePheArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 DB 781 GTTGTCTGCCACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGCGATTCGC 840
 QY 281 AlaAseTyrSerAseProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAse 300
 DB 841 GCTAACTACTCTAAACCCAGCACACGCGCTTCTGTGTGTGCGCACCATCTGAGCAAC 900
 QY 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAsePheMetArgGlnArgIleGlnArg 320
 DB 901 GATCGCTTACGTGCGATTTGGGAACAAGAGCTGATGATGCGCCGCGTATTTCAGCGT 960
 QY 321 MetArgGlnLeuPheValAseThrLeuGlnGlyLysGlyAlaAseArgAsePheSerPhe 340
 DB 961 ATGCTGCTGTTGCTGCTCAATACGCTCGAGAAAGGCGCAACCGGACTTCAGCTTT 1020
 QY 341 IleIleLysGlnAseGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 DB 1021 ATCATCAACAGAACGCGATGTTCTCTCTCAGTGGCTGACAAAGAACAAAGTGTGCGT 1080
 QY 361 LeuArgGluGluPheGlyValTyrAlaValAseSerGlyArgValAseValAlaGlyMet 380
 DB 1081 CTGCGCAAGAGTTTGGCGTATATGCGGTGCTTCTGCTGCGTAAATGTGGCGGATG 1140
 QY 381 ThrProAseAseMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 DB 1141 ACACAGATAACATGGCTCGCTGCGCAAGCGATTTGTCGAGTGTCTG 1188
 RESULT 2
 ADS46042
 ID ADS46042 standard; cDNA; 1191 BP.
 XX
 AC ADS46042;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polynucleotide #785.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polynucleotide; gene; ss.
 XX
 OS Bacteria.
 XX
 PN US20032333675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX

DR WPI; 2004-061375/06.
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 24472; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at: seqdata.uspto.gov/sequence.html.

XX
SQ Sequence 1191 BP; 291 A; 305 C; 317 G; 278 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8,79e-203 Length: 1191
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0

US-10-673-786A-2 (1-396) x ADS46042 (1-1191)

QY 1 MetPheGluAsnIleThrAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
DB 1 ATGTTTGAGAACATTACCGCGCTCTGCGGACCGGATTCCTGGGCTGGCCGATCTGTTT 60
QY 21 ATGAlaAspGluArgProGlyIleAenLeuGlyIleGlyValTyrlsAspGluThr 40
DB 61 CTGGCGGATGAACGTCGCGGCAAAATTAACCTCGGGAATGGTGTCTATTAAGATGAGACG 120
QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrlsLeuGluAsnGlu 60
DB 121 GGCAGAACCCCGGTACTGACGAGCGTGAAGAGGCTGAACATCTGCTCGAATGNA 180
QY 61 ThrThrLysAsnTyrlsLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
DB 181 ACCACCAAAATTAACCTCGGCAATGACGCGATCCCTGAATTTGGTTCGCTGCACTCAGGAA 240
QY 81 LeuLeuPheGlyLysGlySerAlaIleAenAspLysArgAlaArgThrAlaGlnThr 100
DB 241 CTGCTGTTTGGTAAAGGTAGCCCTGTGATCAATGACAAACGTCGTCGACGCGACAGACT 300
QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
DB 301 CCGGGGGGCACTGGGGGCACTACGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 360
QY 121 LysArgValTrrValSerAsnProSerTrrProAsnHisLysSerValPheAsnSerAla 140
DB 361 AAGCGTGTGTGGTGAGCAACCAAGCTGGCGGCAACCAATGACGAGCGTCTTTAACTCTGCA 420

QY 141 GlyLeuGluValArgGluTyrlsAlaTyrlsAspAlaGluAsnHisThrLeuAspPheAsp 160
DB 421 GGTCTGGAGTTTCGTGAATACGCTTATTATGATGCGGAAATCACACTCTTGACTTCGAT 480
QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
DB 481 GCACCTGATTAAACAGGCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTCATGCTGC 540
QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrrPrlnThrLeuAlaGlnLeu 200
DB 541 TGCCATTAACCCCAACCGGTATACGCTTGGAAACAATGGCAACACTTGGCACAACATC 600
QY 201 SerValGluLysGlyTrrLeuProLeuPheAspPheAlaTyrlsGlnGlyPheAlaArgGly 220
DB 601 TCCGTTTGAGAAAGGCTGGTTACCGCTGTTTACGCTTACCGGTTTGGCCGCTGGT 660
QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
DB 661 CTGGAAGAAGATGCTGAAGGACTGGCGGCTTTCGCGGCTATGCATAAAGAGCTGATTGT 720
QY 241 AlaSerSerTyrlsSerLysAsnPheGlyLeuTyrlsAsnGluArgValGlyAlaCysThrLeu 260
DB 721 GCAGTGTCTTACTCTTAAACACTTTGGCCTGTACAACGAGCGTGTGGCCGCTGTACTCTG 780
QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
DB 781 GTTGTGCGGACAGTGAACCGTTCGATCGCGGATTCAGCCAATGAAGCGGCGATTCGC 840
QY 281 AlaAsnTyrlsSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
DB 841 GCTAACTACTCTTAACCCACCAGCACACGCGCGCTTCTGTGTGGCCACCATCTCTGAGCAAC 900
QY 301 AspAlaLeuArgAlaIleTrrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
DB 901 GATGCTTACGTCGCGATTTGGGAAACAAGAGCTGACTGATATGCGCCGCGTATTCAGCGT 960
QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyValAsnArgAspPheSerPhe 340
DB 961 ATGCTCAGTGTGTCGTCATACGCTGAGGAAAGGCGCAACCGCGACTTCAGCTTT 1020
QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
DB 1021 ATCATCAACAGACAGCGCATGTTCTCTTCACTGCGCTGACAAAGAACAGTGTGCTG 1080
QY 361 LeuArgGluGluPheGlyValTyrlsAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
DB 1081 CTGCGCGAAGAGTTTGGCGGTATATCGGCTATATCGGCTGCTTCTGTCGCTAAATGTGCGCGGATG 1140
QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
DB 1141 ACACCAAGATTAACATGCTCCGCTGTCGAGGCAATTTGGCAGTGTCTGCTG 1188
RESULT 3
ADW95403
ID ADW95403 standard; DNA; 1191 BP.
XX
AC ADW95403;
DT 07-APR-2005 (first entry)
XX
DE Nucleotide sequence of an Escherichia coli polypeptide.
XX transgenic plant; glutamic acid dehydrogenase; GDH; ECASPC;
KW 2-oxo glutaric acid; agriculture; gene; ds.
XX Escherichia coli.
OS
FH Key Location/Qualifiers
FT CDS 1..1191
FT /tag= a
XX
PN WO200506847-A1.
XX

PD 27-JAN-2005.
 XX 15-JUL-2004; 2004WO-JP010451.
 XX 17-JUL-2003; 2003JP-00198559.
 XX (AJIN) AJINOMOTO CO INC.
 XX Kisaaka H, Miwa T, Akiyama A;
 XX WPI; 2005-132242/14.
 XX P-PSDB; ADW95404.
 XX Producing plant having improved growth and yield under cultivation
 PT conditions with decrease in nitrogen, involves expressing transduced
 PT glutamic acid dehydrogenase gene in plant so as to increase their 2-oxo
 PT glutaric acid contents.
 XX Disclosure; SEQ ID NO 19; 94pp; Japanese.
 XX The specification describes a method of producing a plant having improved
 CC growth and yield under cultivation conditions with a decrease in
 CC nitrogen. The method involves transducing glutamic acid dehydrogenase
 CC (GDH) or ECASPC genes into a plant and expressing the gene so as to
 CC increase 2-oxo glutaric acid content of the plant, or applying proline to
 CC the foliage of the plant so as to increase the 2-oxo glutaric acid
 CC content of the plant. The method is useful for producing a plant having
 CC improved growth and yield under cultivation conditions with a decrease in
 CC nitrogen. The present sequence encodes a polypeptide, used in the course
 CC of the invention.
 XX
 XX SQ Sequence 1191 BP; 291 A; 305 C; 317 G; 278 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 8,79e-203 Length: 1191
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 14 Gaps: 0
 US-10-673-786A-2 (1-396) x ADW95403 (1-1191)
 Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 1 ATGTTTGAGAACATTACCGCGCTCTCCGACCGGATTCTGGGCTGGCGGATCTGTTT 60
 Qy 21 ArgAlaaspGluArgProGlyIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 Db 61 CGTCCGATGAACGTCCCGGCAAAATTAACCTCGGGATTGGTGTCTATAAAGATGAGACG 120
 Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAAsnGlu 60
 Db 121 GGCAAAACCCGGTACTGACGAGCGTGAAAGGCTGAAAGGCTGACAGATATCTGCTCGAANAATGAA 180
 Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 181 ACCACCAAAATTAACCTCGGATTCGACGCGATCCCTGAATTTGGTCTGCTGACCTCAGGAA 240
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 241 CTGCTGTTTGGTAAAGTAGCGCTGATCAATGACAAACGTGCTCGCACGGCACAGACT 300
 Qy 101 ProGlyGlyThrGlyValAlaArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 301 CCGGGGGGACCTGGCGCACTAGCGGTGCTCCGATTTCTGGCAAAAATTAACCGCGTT 360
 Qy 121 LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 Db 361 AAGCGTGTGGTGAGCAACCCAGCTGGCGCAACCATTAAGAGCGTCTTTAACTCTGCA 420
 Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAAsnHisThrLeuAspPheAsp 160

Db 421 GGTCGGAAGTTCTGTAATACGCTTATTATGATGCGGAAATACACACTCTTGTGACTTCGAT 480
 Qy 161 AlaLeuIleAsnSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 481 GCACCTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTCTTCCATGGCTGC 540
 Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
 Db 541 TGCATAAACCCACCGGTATCGACCTTACGCTGACCAATAGGCAACACTGGCACAACTC 600
 Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 601 TCGTTGAGAAAGGCTGTTTACCGCTGTTGACTTCGCTTACCAAGGTTTTTGGCCGTGT 660
 Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 Db 661 CTGGAAGAAGATGCTGAAGGACTGCGGCTTTCGCGCTATGCATAAAGAGCTGATTGTT 720
 Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 Db 721 GCCAGTTCTTACTCTAAAAAATTTGGCTGTGTACCAACGAGCGTGTGGCGCTTGTACTCTG 780
 Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 781 GTTGTGCGCACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGCGCATTCGC 840
 Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 Db 841 GCTAACTACTCTAACCCACCAGCACACGGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 900
 Qy 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnAtgIleGlnArg 320
 Db 901 GATCGGTTTACGTCGATTGGGAACAAGAGCTGATATGCGCCAGCGATTTCAGCGT 960
 Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
 Db 961 ATGGCTCAGTTGTCTGTCATACGCTCAGGAAAGGCGCAACCGGACTTCAGCTTT 1020
 Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1021 ATCATCAACAGAACGCGATGTTCTCTTCTCAGTGGCTGACAAAAGAACCAAGTGTGCGT 1080
 Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1081 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGTAATATGTCGCGGATG 1140
 Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 1141 ACACCAATTAACATGGCTCGCTGTGGAAGCGATTGTGGCAGTGTCTG 1188
 RESULT 4
 AAV40259
 ID AAV40259 standard; DNA; 1331 BP.
 XX
 AC AAV40259;
 XX
 DT 13-OCT-1998 (first entry)
 XX
 DE Escherichia coli aspC gene.
 XX
 KW Brevibacterium lactofermentum; lysC; L-lysine; coryneform bacterium;
 KW aspartokinase; feedback inhibition; dihydropicolinate reductase;
 KW diaminopimelate decarboxylase; aspartate aminotransferase; ds.
 XX
 OS Escherichia coli.
 XX
 FH Location/Qualifiers
 CDS 10..1197
 FT /*tag= a
 FT /product= "aspC"
 XX
 PN EP854189-A2.
 XX

PD 22-JUL-1998.
 XX
 XX 05-DEC-1997; 97EP-00121443.
 XX
 XX 05-DEC-1996; 96JP-00325659.
 XX
 XX (AJIN) AJINOMOTO CO INC.
 XX
 XX Araki M, Sugimoto M, Yoshihara Y, Nakamatsu T;
 XX
 XX WPI; 1998-379060/33.
 XX
 XX P-PSDB; AAW69553.
 XX
 XX Recombinant DNA autonomously replicable in coryneform bacteria - used to
 XX produce L-lysine, codes for e.g. aspartokinase, di:hydronicotinase,
 XX reductase and synthase and di:amino-pimelate decarboxylase.
 XX
 XX Example 5; Page 37-38; 59pp; English.
 XX
 XX The present invention describes a recombinant DNA autonomously replicable
 XX in cells of coryneform bacteria (CB), comprising a DNA sequence coding
 XX for an aspartokinase (AK) in which feedback inhibition by L-lysine and L-
 XX threonine is desensitized, a DNA sequence coding for a
 XX dihydronicotinase reductase (DHPR), a DNA sequence coding for
 XX diaminopicolinate synthase (DHPS), a DNA sequence coding for
 XX diaminopimelate decarboxylase (DAMD) and a DNA sequence coding for
 XX aspartate aminotransferase (AAT). The present sequence encodes aspC from
 XX Escherichia coli. The DNA and related products from the present
 XX invention, can be used for improving L-lysine productivity by CB. The L-
 XX lysine produced can be used as a fodder additive
 XX
 XX SQ Sequence 1331 BP; 330 A; 340 C; 350 G; 311 T; 0 U; 0 Other;
 XX
 XX Alignment Scores:
 XX Pred. No.: 1.03e-202 Length: 1331
 XX Score: 2045.00 Matches: 396
 XX Percent Similarity: 100.0% Conservative: 0
 XX Best Local Similarity: 100.0% Mismatches: 0
 XX Query Match: 100.0% Indels: 0
 XX DB: 2 Gaps: 0
 XX
 XX US-10-673-786A-2 (1-396) x AAV40259 (1-1331)
 QY 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 DB 10 ATGTTTGAGAACATTACCGCGCTCTGCGCCGACCGGATTCCTGGCGCTGGCGATCTGTTT 69
 QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyLysIleGlyValTyrLysAspGluThr 40
 DB 70 CGTCCGATGAACGTCCCGGCAAAATTAACCTCGGATTGGTGTCTATAAAGATGAGACG 129
 QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
 DB 130 GGCAGAAACCCCGGTACTGACACGCGTGAAGGCTGAACAGTATCTGCTCGAAATGAA 189
 QY 61 ThrThrLysAsnTyrLeuGlyLysIleAspGlyLysIleProGluPheGlyArgCysThrGlnGlu 80
 DB 190 ACCACCAAAATTAATTCCTCGGATTTGACCGCATCTCTGAAATTTGGTGGCTGACCTCAGGAA 249
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 DB 250 CTGCTGTTTGTAAAGGTAGCGCTGATCAATGACAAACGCTGTCGACGCGACAGACT 309
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 DB 310 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCCGATTTCTGGCAAAAATATACAGCGTT 369
 QY 121 LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 DB 370 AAGCGTGTGGGTGAGCAACCCAGCTGGCGCAACCAATAGAGCGCTTTTAATCTTGCA 429
 QY 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAspPheAsp 160

DB 430 GGTCTGGAAGTTCGTGAATAGCTTATTATGATCGGAAATACACACTCTTGACTTCGAT 489
 QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 DB 490 GCCTGATTAAACAGCTGAATGAAGCTCAGCTGCGAGCGTAGTGTGTTCCATGGCTGC 549
 QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
 DB 550 TGCCATAACCAACCGGTATCGACCTACGCTGGAACAATGCAACAACTGGCAACATC 609
 QY 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 DB 610 TCCGTTTGAGAAAGCTGTTTACCGCTGTTTACCTTCGCTTACCAAGGTTTTCGCCGTGT 669
 QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
 DB 670 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCGCGCTATGCATAAAGAGCTGATTGTT 729
 QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 DB 730 GCCAGTTCCTACTCTAAAAACCTTTGGCTGTGTACACAGCGTGTGGCGCTTGACTCTG 789
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 DB 790 GTTGCTCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGCGGATTCGC 849
 QY 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 DB 850 GCTNACTACTCTAACCCACGACGACGCGCTTCTGTGTTGCCACCATCTCGAGCAAC 909
 QY 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 DB 910 GATCGGTTACGTGCGATTTGGGAACAAGAGCTGATGATGCGCCAGCGTATTTCAGCGT 969
 QY 321 MetArgGlnLeuPheValAsnThrLeuGluLysGlyAlaAsnArgAspPheSerPhe 340
 DB 970 ATGCGTCAGTGTGTTGCGTCAATACCTCGAGAAAGCGCAACCCGCGACTTCAGCTTT 1029
 QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 DB 1030 ATCATCAACAGAACGCGCATGTTCTCTTCAGTGGCTGACAAAGAACAAAGTCTGCGT 1089
 QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 DB 1090 CTGCGCGAAGAGTTTGGCTGATATGCGGTTCTTCTGCTGCGTAAATGTGCGCGGATG 1149
 QY 381 ThrProAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 DB 1150 ACACAGATAACATGGCTCCGCTGTGCGAAGCGATTGTGGCAGTGCTG 1197
 XX
 XX RESULT 5
 XX AAN71109
 ID AAN71109 standard; DNA; 3659 BP.
 XX
 XX AC AAN71109;
 XX
 XX 01-JAN-1980 (first entry)
 DT
 XX
 DE pheA aroF aspC operon in plasmid pME219.
 XX
 XX alpha-amylase; feedback inhibition; amino acid synthesis;
 XX composite plasmid; ss.
 KW
 XX
 OS Bacillus licheniformis.
 XX
 XX Key Location/Qualifiers
 XX CDS 69..1241
 XX FT /*tag= a
 XX FT 1259..2329
 XX FT /*tag= b
 XX FT 2344..3534
 XX FT /*tag= c
 XX

PN W08700202-A.
 XX 15-JAN-1987.
 XX 24-JUN-1986; 86WO-US001353.
 XX 24-JUN-1985; 85US-00747732.
 XX (NUTR-) NUTRASWEET CO.
 XX Edwards MR, Taylor PP, Hunter MG, Fotheringh IG;
 XX WPI; 1987-021998/03.
 DR P-PSDB; AAP70752, AAP71677, AAP71678.
 XX
 XX Composite plasmids contg. multiple genes in transcriptional units -
 PT useful for prodn. of amino acid(s), esp. L-phenylalanine and L-tyrosine.
 XX Disclosure; Page 38; 57pp; English.
 XX
 XX This sequence may be inserted into a composite plasmid and used for the
 CC production of amino acids. See also AAN71053-55, AAN71107, AAN71109,
 CC AAN71111 and AAP70696-97 and AAP70750, AAP70752 and AAP70754
 XX
 XX Sequence 3659 BP; 936 A; 911 C; 947 G; 865 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 4.19e-202 Length: 3659
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservativeness: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 1 Gaps: 0
 US-10-673-786A-2 (1-396) x AAN71109 (1-3659)
 Qy 1 MetPheGluAenIleThrAlaAProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
 Db 2344 ATGTTTGAGACATTACCGCGCTCTGCCGACCGATTCTGGCGCTGGCGGATCTGTTT 2403
 Qy 21 ArgAlaAaspGluArgProGlyLysIleAenLeuGlyLysValThrLysAaspGluThr 40
 Db 2404 CGTCCGATGAACCTCCCGGAAAAATTAACCTCGGGATTGGTGTCTATAAAGATGAGACG 2463
 Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
 Db 2464 GGCAAAACCCCGTACTGACCGCGTGGAAAGGCTGAACAGTATCTGCTCGAAATGAA 2523
 Qy 61 ThrThrLysAenTyrLeuGlyLysIleAaspGlyLysProGluPheGlyArgCysThrGlnGlu 80
 Db 2524 ACCACCAAAATTAACCTCGGCATTCAGCGCATCCCTGAATTTGGTGTGCTGACCTCAGGAA 2583
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAaspLysArgAlaArgThrAlaGlnThr 100
 Db 2584 CTGCTGTTTGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGACGCGCACAGCT 2643
 Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAaspPheLeuAlaLysAenThrSerVal 120
 Db 2644 CCGGGGGGCACTGGCGGCACTACGGGTGCTGCCGATTTCTTGGCAAAAAATACCGCGTT 2703
 Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
 Db 2704 AAGCGTGTGGTGAGCAACCCCAAGCTGGCGGACCATTAAGAGCGCTTTTAACCTGCA 2763
 Qy 141 GlyLeuGluValArgGluTyrAlaTyrThrAaspAlaGluAenHisThrLeuAaspPheAasp 160
 Db 2764 GGTCTGGAAGTTCGTGAATACGCTTATTATGATCGGAAAAATCACACTCTTGACTTCGAT 2823
 Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCys 180
 Db 2824 GCACGTGATTAAACGCCGTAATGAAGCTCAGGCTCGGACGAGTGTGCTTCCATGGCTGC 2883
 Qy 181 CysHisAenProThrGlyIleAaspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200

Db 2884 TGGCATAAACCCACCGGTATCGACCTTACGACCAATGGGCAACACTGGGCAACTC 2943
 Qy 201 SerValGluLysGlyTrpLeuProLeuPheAaspPheAlaTyrGlnGlyPheAlaAargGly 220
 Db 2944 TCGTTTGAGAAAGCTGTTTACCGCTGTTTACCTTCGTTTACCGAGGTTTGGCCGTTGT 3003
 Qy 221 LeuGluGluAaspAlaGluGlyLeuArgAlaPheAlaAalaMethHisLysGluLeuIleVal 240
 Db 3004 CTGGAAGACATGCTGAAGGACTGCGGCTTTCGCGCTATGCATAAAGAGCTGATTGTT 3063
 Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyAlaCysThrLeu 260
 Db 3064 GCCAGTTCTTACTCTAAAAACTTTGGCCTGTACACAGCGGTGTGGCGCTTGTACTCTG 3123
 Qy 261 ValAlaAaspSerGluThrValAaspArgAlaPheSerGlnMetLysAlaAalaIleArg 280
 Db 3124 GTTGCTGCCACAGTGAACCGTTGATCGGCGATTTCAGCCAAATGAAGGCGGATTCGC 3183
 Qy 281 AlaAenTyrSerAenProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
 Db 3184 GCTAACTACTCTTAACCCACGACACGCGGCTTCTGTTGTTGCCACCATCTCAGCAAC 3243
 Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAaspMetArgGlnArgIleGlnArg 320
 Db 3244 GATCGCTTACGTGCGATTTGGGAACCAAGAGCTGACGTATATGCGCCAGCGTATTCAGCGT 3303
 Qy 321 MetArgGlnLeuPheValAenThrLeuGlnGluLysGlyAlaAenArgAaspPheSerPhe 340
 Db 3304 ATCGCTCAGTTGTTGTTCAATACGCTGCAGGAAAGGCGCAAAACCGCGACTTCAGCTTT 3363
 Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 3364 ATCATCAACAGAACGCGATGTTCTCTTTCAGTGGCTGACAAAGAAACAAGTGTGCGT 3423
 Qy 361 LeuArgGluPheGlyValTyrAlaValAlaSerGlyArgValAenValAlaGlyMet 380
 Db 3424 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGCGTAAATGTGGCGCGGATG 3483
 Qy 381 ThrProAaspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 3484 ACACGATTAACATGGCTCCGCTGTGCGAAGCGATTGTGCGAGTGTCTG 3531
 RESULT 6
 AAS46273/c
 ID AAS46273 standard; DNA; 14759 BP.
 XX
 AC AAS46273;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE DNA encoding novel mar regulated protein (NIMR) #42.
 XX
 XX mar regulated polypeptide; NIMR; microbial infection; antibacterial; ds.
 OS Escherichia coli.
 OS
 PN WO200170776-A2.
 XX
 PD 27-SEP-2001.
 XX
 XX 08-MAR-2001; 2001WO-US007478.
 XX
 PR 10-MAR-2000; 2000US-0188362P.
 XX
 PA (TUFT) TUFTS COLLEGE.
 XX
 PI Levy SB, Barbosa TM, Alekshun MN;
 XX
 DR WPI; 2001-602769/68.
 DR P-PSDB; AAU29374.
 XX
 PT Identifying compounds that modulate a newly identified mar regulated

PT polypeptide activity, useful as antimicrobial compounds, involves
 XX contacting the polypeptide with a test compound.
 PS Disclosure; Page 477-485; 526pp; English.

XX The invention relates to a method of identifying compounds that modulate
 CC a newly identified mar regulated (NMR) polypeptide activity. The method
 CC comprises contacting an NMR polypeptide with a test compound under
 CC interaction conditions, determining the ability of the compound to
 CC modulate the activity or expression of the polypeptide, and selecting the
 CC modulators. NMR nucleic acids and polypeptides are used in the treatment
 CC of microbial infections, and in screening for modulators of NMR
 CC expression and activity. These modulators can be used to reduce the
 CC infectivity of a microbe on a surface, and the virulence of a microbe in
 CC a subject suffering from an infection. AAS46273-AAS46278 represent
 CC Escherichia coli NMR coding sequences of the invention

SQ Sequence 14759 BP; 3703 A; 3840 C; 3713 G; 3503 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2.92e-201 Length: 14759
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0

US-10-673-786A-2 (1-396) x AAS46273 (1-14759)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 2755 ATGTTTGAACAATACCGCGCTCTCGCCAGCCGATTCTGGCGCTGGCGATCTGTTT 2696
 Qy 21 ArgAlaAspGluAcpProGlyIleAsnLeuGlyIleGlyValTyrLeuAspGluThr 40
 Db 2695 CGTCCGATGAACGTCGCGGCAAAATTAACCTCGGGATTGGTGTCTATAAGATGAGACG 2636
 Qy 41 GlyIysThrProValLeuThrSerValIysValAlaGluGlnTyrLeuGluAenGlu 60
 Db 2635 GGCACAAACCCGGTACTACCGACGCTGAAAGGCTGACAGTATCTCTCGAAATGAA 2576
 Qy 61 ThrThrIysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 2575 ACCACAAAATTAACCTCGGATTCGCGCATCCCTGAATTTGGTCGCTGCACACAGAA 2516
 Qy 81 LeuLeuPheGlyIysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 2515 CTGCTGTTTGGTAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCACGACAGACT 2456
 Qy 101 ProGlyGlyThrGlyValAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 2455 CCGGGGGGCACTGGCGCACTACCGGTGCTGCCGATTTCTGGGCAAAAATAACACGCTT 2396
 Qy 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 Db 2395 AAGCGTGTGGTGAGCAACCCAGCTGGCGAACCAATAGAGCGTCTTTAACTCTGCA 2336
 Qy 141 GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAsnHisThrLeuAspPheAsp 160
 Db 2335 GGTCTGGAAGTTCTGGAATACGCTTATTATGATGCGGAAATCAACACTCTTGTACTTCGAT 2276
 Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 2275 GCACGTGATTACAGCCTGAAAGTGAAGCTCAGGCTGGCGACGTAGTCTGTTCCATGGCTGC 2216
 Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
 Db 2215 TGCCATAACCAACCGGTATCGACCTACGCTGGAGCAATGGCAAACTGGCACAACTC 2156
 Qy 201 SerValGluIysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 2155 TCCGTTGAGAAAGGCTGTTACCGCTGTTTGCATTCGCTTACCGAGGTTTTCGCCGCTGTT 2096

Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 Db 2095 CTGGAAGAAGATGCTGAAGGACTCGCGCTTTCGGCGCTATGCATAAAGAGCTGATGTTT 2036
 Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyValaCysThrLeu 260
 Db 2035 GCCAGTTTCTTACTCTAAANAATTTGGCTGTACAAACGAGCGTGTGGCGCTTGTACTCTG 1976
 Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 1975 GTTGCTGCCGACAGTGAACCCGTTGATCGCGCATTCAGCCAAATGAAAGCGCGATTGCG 1916
 Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValAlaAlaThrIleLeuSerAsn 300
 Db 1915 GCTAACTACTCTAACCCACACGACACGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 1856
 Qy 301 AspAlaLeuArgAlaIleTyrGluGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 1855 GATCGTTTACGTGCGATTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTGAGCGT 1796
 Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGlyLysGlyAlaAsnArgAspPheSerPhe 340
 Db 1795 ATGCGTCAGTTGTTCTCAATACGCTGCAGGAAAAGGCGCAACCGCGACTTTCAGCTTT 1736
 Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1735 ATCATCAACAGAACGCGCATGTTCTCTTCAGTGGCTTGACAAAAGAACAGTGTGCTG 1676
 Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1675 CTGGCGGAAGAGTTGGCGTATATGCGGTTCCTCTGCTCGCGTAAATGTGGCCGGGATG 1616
 Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 1615 ACACAGATAACATGCTGCTCGCTGTGCAAGCGATTGTGGCAGTGTCTG 1568

RESULT 7
 ID AAN71108
 ID AAN71108 standard; DNA; 1293 BP.
 AC AAN71108;
 XX
 DT 01-JAN-1980 (first entry)
 XX
 DE Optimized Escherichia coli aspC gene.
 XX
 KW aspC gene; feedback inhibition; amino acid synthesis; composite plasmid;
 KW ss.
 XX
 OS Escherichia coli.
 XX
 PN W08700202-A.
 XX
 PD 15-JAN-1987.
 XX
 PF 24-JUN-1986; 86WO-US001353.
 XX
 PR 24-JUN-1985; 85US-00747732.
 XX
 PA (NUTR-) NUTRASWEET CO.
 XX
 PI Edwards MR, Taylor PP, Hunter MG, Fotheringh IG;
 XX
 DR WPI; 1987-021998/03.
 DR P-ESDB; AAP70751.
 XX
 PT Composite plasmids contg. multiple genes in transcriptional units -
 XX useful for prodn. of aminoacid(s), esp. L-phenylalanine and L-tyrosine.
 PS Disclosure; Page 25; 57pp; English.
 XX
 CC This sequence may be inserted into a composite plasmid and used for the
 CC production of amino acids. See also AAN71053-55, AAN71107, AAN71109-11

5

[illegible]

```
Pred. No.: 5,45e-165 Length: 1191
Score: 1682.00 Matches: 316
Percent Similarity: 89.4% Conservative: 38
Best Local Similarity: 79.8% Mismatches: 42
Query Match: 82.2% Indels: 0
DB: 10 Gaps: 0

US-10-673-786A-2 (1-396) x ACF68649 (1-1191)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProfileLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTTTGAGAAATCATCCGAGCGCTGCCGACCTATTCTTGGCTTAGCCGATGTTTC 60
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 61 CGTTCTGATCTCTACAAATAAAATCAACCTTGGTATCGGTGTCTATAAAGACGAACA 120
Qy 41 GlyLysThrProValLeuThrSerValLysIleAlaGluGlnTyrLeuGluAsnGlu 60
Db 121 GGAATAACCCCGGTTCTGACGAGTGTAAAAAGCTGAACAATATTTACTGGAAAAACGA 180
Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 181 ACAACAAGAAATATCTGCCGATTAGCGGCTAGCCGAATTTGGCCGCTAACTCAAGAA 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 241 TTACTGTTTGGCAAGATACCCAGTGTGTACAGATAAACGCGCCGACAGCACAAAGC 300
Qy 101 ProGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 301 CCAGGCGGTACCGGTGCTTACGATTTGCTGCGGATTTTATTCGCAACAGACTAAATGCT 360
Qy 121 LysArgValTyrPheSerAsnProSerTyrProAsnHisLysSerValPheAsnSerAla 140
Db 361 AAACGAGTTGGATCAGCAACCAACCTGGCCAAACCATATAAAGCTTTTTCGCGCGCT 420
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 421 GGTCTGGAAGTCTGTGAATATAAATACTATGACGCTGAAACCAACGCGCTGAATTCGAA 480
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 481 GACATGTGCGCAACGCTGTCGGAAGCTCAGCTGCTGATGTGTCTTCCACGCGCTGC 540
Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 541 TGCCACAAATCCGACAGGATCGATCCAAACCCGCGCAATGGGCTAACTGGCAGAAATG 600
Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 601 TCTCGGAGAAAGCTGGTTGGCTATTTTGGTATTCGCTTACCAAGGATTTGCCAAGGC 660
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 661 CTAAACGAAGATGACAGAGGCTACGATTTTTCGGAATAATCATATGAAGTATGATGTT 720
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 721 GCCAGCTCTTACTCCAAANAACTTTGGCTGTACAAATGAACGTGTCGCTGCTACTATT 780
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 781 GTTCTAGTGACAGTATACAGCAAGAAAGCGTTACGCAAGCTAAAGCGATTTATCCGT 840
Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 841 GCTAACTATTCCAAACCCACCGCTCATGGTGCATCTATTGTCTACATACATTTTGTCAAAT 900
Qy 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 901 GAAGACCTGAAGACGCTTGGGAACAGAACTGACCACTGACCACTGCGCGAACGATTCAGCGT 960
```

```
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 961 ATCGCTCAACTGTTTGTCAATACCTTGCAGAAAAAGGGCAAAACAGGATTTTCAGCTTT 1020
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1021 ATTATTAGCCAAATGGTATGTTCTCATTCAGTGGCTGACAAAAAGAACAGTAGAACGT 1080
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1081 CTGCGTGATGAGTTGTTGTTATATACGCTGTTCAGTTCGCTCGTATTAACTGCTGGCTTG 1140
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1141 ACGTTAGAAAAACATGGCTCCACTATGTGAAGCCATTGTTGCAGTACTC 1188

RESULT 10
ACF67367_14
Continuation (15 of 57) of ACF67367 from base 1400001 (Photorhabdus luminescens nucleoti
WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367
WP Fragment Name Begin End
WP ACF67367_00 1 110000
WP ACF67367_01 100001 210000
WP ACF67367_02 200001 310000
WP ACF67367_03 300001 410000
WP ACF67367_04 400001 510000
WP ACF67367_05 500001 610000
WP ACF67367_06 600001 710000
WP ACF67367_07 700001 810000
WP ACF67367_08 800001 910000
WP ACF67367_09 900001 1010000
WP ACF67367_10 1000001 1110000
WP ACF67367_11 1100001 1210000
WP ACF67367_12 1200001 1310000
WP ACF67367_13 1300001 1410000
WP ACF67367_14 1400001 1510000
WP ACF67367_15 1500001 1610000
WP ACF67367_16 1600001 1710000
WP ACF67367_17 1700001 1810000
WP ACF67367_18 1800001 1910000
WP ACF67367_19 1900001 2010000
WP ACF67367_20 2000001 2110000
WP ACF67367_21 2100001 2210000
WP ACF67367_22 2200001 2310000
WP ACF67367_23 2300001 2410000
WP ACF67367_24 2400001 2510000
WP ACF67367_25 2500001 2610000
WP ACF67367_26 2600001 2710000
WP ACF67367_27 2700001 2810000
WP ACF67367_28 2800001 2910000
WP ACF67367_29 2900001 3010000
WP ACF67367_30 3000001 3110000
WP ACF67367_31 3100001 3210000
WP ACF67367_32 3200001 3310000
WP ACF67367_33 3300001 3410000
WP ACF67367_34 3400001 3510000
WP ACF67367_35 3500001 3610000
WP ACF67367_36 3600001 3710000
WP ACF67367_37 3700001 3810000
WP ACF67367_38 3800001 3910000
WP ACF67367_39 3900001 4010000
WP ACF67367_40 4000001 4110000
WP ACF67367_41 4100001 4210000
WP ACF67367_42 4200001 4310000
WP ACF67367_43 4300001 4410000
WP ACF67367_44 4400001 4510000
WP ACF67367_45 4500001 4610000
WP ACF67367_46 4600001 4710000
WP ACF67367_47 4700001 4810000
WP ACF67367_48 4800001 4910000
WP ACF67367_49 4900001 5010000
WP ACF67367_50 5000001 5110000
WP ACF67367_51 5100001 5210000
```


WP	ACF67367_52	5200001	5310000
WP	ACF67367_53	5300001	5410000
WP	ACF67367_54	5400001	5510000
WP	ACF67367_55	5500001	5610000
WP	ACF67367_56	5600001	5648994

Alignment Scores:

Pred. No.: 2,97e-162 Length: 110000

Score: 1682.00 Matches: 316

Percent Similarity: 89.4% Conservative: 38

Best Local Similarity: 79.8% Mismatches: 42

Query Match: 82.2% Indels: 0

DB: 10 Gaps: 0

US-10-673-786A-2 (1-396) x ACF67367_14 (1-110000)

Qy	1	MetPheGluAsnIleThrAlaLaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe	20
<hr/>			
Dd	23013	ATGTTTGAGAAATCACCGAGCGCCTTCGGACCCTATTCTTGGCTTAGCCGATAGTTTC	23072
<hr/>			
Qy	21	ArgAlaAaspGluArgProGlyLyseIleAsnLeuGlyIleGlyValTyrllysAspGluThr	40
<hr/>			
Dd	23073	CGTTCGTGACTCGTACAAAATAAATCAACCTTGGTGATCGGTGTCTATATAAGACGAACA	23132
<hr/>			
Qy	41	GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrlzeuLeuGluAsnGlu	60
<hr/>			
Dd	23133	GGAAAAACCCGGTCTCTGACAGTGTAAAAAAGCTGAACAATATTACTTGGAAAAACGAA	23192
<hr/>			
Qy	61	ThrThrLysAsnTyrlzeuGlyIleAaspGlyIleProGluPheGlyVargCyethrGlnGlu	80
<hr/>			
Dd	23193	ACAAAGAAGTAATTATCTGCCGATTACGGCTTAGCCGAATTTTGGCCGCCTAACTCAAGAA	23252
<hr/>			
Qy	81	LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaAargThrAlaGlnThr	100
<hr/>			
Dd	23253	TTACTGTTGGCAAAGATCACCCAGTTGTGCACAGATAAACGCCCGCACACACCAAGC	23312
<hr/>			
Qy	101	ProGlyGlyThrGlyAlaLeuArgValAlaAlaAaspPheLeuAlalysAsnThrSerVal	120
<hr/>			
Dd	23313	CCAGGCGGTACCCGGTGCTTTACGTTTTTGTGTCGGATTTCAITGCCAAAACAGACTAATGT	23372
<hr/>			
Qy	121	LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla	140
<hr/>			
Dd	23373	AACAGAGTTTGGATCAGCAACCCACCTGCCAAACCATAAANAACGTTTTTTCCCGCGCT	23432
<hr/>			
Qy	141	GlyLeuGluValArgGluTyAlaTyTrAspAlaGluAsnHisThrLeuAaspPheAsp	160
<hr/>			
Dd	23433	GGTCTGGAAGTCTGTGNAATAAATACTATGACGCTGAAAAACACCGCGCTGAATTCGAA	23492
<hr/>			
Qy	161	AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAaspValLeuPheHisGlyCys	180
<hr/>			
Dd	23493	GACATGTGCGAAGCCTGTCGGAAGCTCAGGCTGGTGATGTTGTTCTGTTCCACGCTGC	23552
<hr/>			
Qy	181	CysHisAsnProThrGlyIleAaspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu	200
<hr/>			
Dd	23553	TGCCACAATCCGACGACATCGATCCAACCCCAGCAATGGCTTAACTGGCGCAAAATG	23612
<hr/>			
Qy	201	SerValGlnLysGlyTrpLeuProLeuPheAaspPheAlaTyTrGlnGlyPheAlaArgGly	220
<hr/>			
Dd	23613	TCCTCGGAGAAAGCGTGGTGGCTATTATTTTGTATTCGCTTACCAAGGATTTGCCAAGGCG	23672
<hr/>			
Qy	221	LeuGluGluAaspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal	240
<hr/>			
Dd	23673	CTAAACGAAGATGCAGAGGCGCTACGTATTTTTTGGCGAAAAATCATATGAACGTATGTT	23732
<hr/>			
Qy	241	AlaSerSerTyrlserLysAsnPheGlyLeutyTrAsnGluArgValGlyAlaCysThrLeu	260
<hr/>			
Dd	23733	GCCAGCTCTTACTCCAAAACTTTGGCCGTGTACAATGAACGTGTCCGTCGCTGTACTATT	23792
<hr/>			
Qy	261	ValAlaAlaAaspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg	280
<hr/>			
Dd	23793	GTTCGTAGTGACAGTAGTACGCAAGAAAGACCGTTCCAGCAAGCTTAACCGATTATTCGT	23852
<hr/>			
Qy	281	AlaAsnTyrlserAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn	300

CC recombinant production of the proteins, particularly toxins and
 CC antibiotics useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens genes
 XX
 SQ Sequence 249878 BP; 75946 A; 56652 C; 49249 G; 68029 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 9.3e-162 Length: 249878
 Score: 1682.00 Matches: 316
 Percent Similarity: 89.4% Conservative: 38
 Best Local Similarity: 79.8% Mismatches: 42
 Query Match: 82.2% Indels: 0
 DB: 10 Gaps: 0

US-10-673-786A-2 (1-396) x ACF65381 (1-249878)

Qy 1 MetPheGluAenIleThrAlaAProAlaAsePProIleLeuGlyLeuAlaAsePLeuPhe 20
 Db 1334 ATGTTTGAGAAATACCGCAGCGCTGCCACCTATTCTTGGCTTAGCGGATGTTTC 1393
 Qy 21 ArgAlaAsePGluAArgProGlyLysIleAenLeuGlyIleGlyValTyrLysAsePLeuThr 40
 Db 1394 CGTCTGATCTCGTACAAATATAATCAACCTTGGTATCGGTGTCTAATAGACGAACA 1453
 Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAseGlu 60
 Db 1454 GGAATAACCCCGGTTCTGACCAAGTGTAAAAAGCTGAACAATATTTACTGGAACGAA 1513
 Qy 61 ThrThrLysAenTyrLeuGlyIleAsePGlyLysProGluPheGlyArgCysThrGlnGlu 80
 Db 1514 ACAACAAAGAAATATTATGCGGCTTAGCGGCTTACCGCAATTTGGCGGTAACTCAAGAA 1573
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAsePysArgAlaAArgThrAlaGlnThr 100
 Db 1574 TTACTGTTTGCAAGATCACCCAGTGTGTACAGTAACCGCGCCGACAGCAACAGC 1633
 Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAsePheLeuAlaLysAenThrSerVal 120
 Db 1634 CCAGCGGTACCGGTGCTTTACGTATTGCTGCGGATTTCAATGTCACAAACAGACTAATGCT 1693
 Qy 121 LysArgValTropValSerAseProSerTropAenHisLysSerValPheAseSerAla 140
 Db 1694 AACCGAGTTTGGATCAGCAACCAACCTGGCCAAACCAATAAAACGTTTTTTCGCGCGCT 1753
 Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAsePAlaGluAenHisThrLeuAsePheAseP 160
 Db 1754 GGTCTGGAGTCTGTGAATATAAATATATATGATGATGATGATGATGATGATGATGATG 1813
 Qy 161 AlaLeuIleAenSerLeuAseGluAlaGlnAlaGlyAsePValValLeuPheHisGlyCys 180
 Db 1814 GACATGCTGGCAAGCGCTGCGAAGCTCAGGCTGCTGATGTTGTTTCTTCCACGGCTGC 1873
 Qy 181 CysHisAenProThrGlyIleAseProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
 Db 1874 TGCCACAATCCGACAGGATCGATCCACCCCGGCACCAATGGGCTTAACCTGGCAGAAATG 1933
 Qy 201 SerValGluLysGlyTropLeuProLeuPheAsePheAlaTyrGlnGlyPheAlaAArgGly 220
 Db 1934 TCTCGGAGAAAGGCTGCTGCTTATTTGATTTTGGCTTACCAAGGATTTGCCAAGGCG 1993
 Qy 221 LeuGluGluAsePAlaGluGlyLeuArgAlaPheAlaAseHisLysGluLeuIleVal 240
 Db 1994 CTAACCAAGATGACAGAGGCTACGTATTTTGGCAAAATATCAATGAATGAATGATGTT 2053
 Qy 241 AlaSerSerTyrSerLysAsePheGlyLeuTyrAseGluArgValGlyAlaCysThrLeu 260
 Db 2054 GCCAGCTTCTACTCCAAAACCTTTGGCGCTGTACAAATGAACGTGCGGTGCTGTACTATT 2113

Qy 261 ValAlaAlaAsePserGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 2114 GTTGTAGTACAGTGTATATATATATATATATATATATATATATATATATATATATAT 2173
 Qy 281 AlaAseTyrSerAseProAlaHisGlyAlaSerValValAlaThrIleLeuSerAse 300
 Db 2174 GCTAACTATTCCAAACCCCGCTCATGGTGCATCTATTGTCTACTACATTTTGTCAAT 2233
 Qy 301 AspAlaLeuArgAlaIleTropGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 2234 GAAGACCTGAAGACAGCTTGGGAACAGGAACCTGACCACTGCGCAACGATATCCAGCT 2293
 Qy 321 MetArgGlnLeuPheValAseThrLeuGlnGlyLysGlyAlaAseArgAsePheSerPhe 340
 Db 2294 ATGCGTCAACTGTTGTCAATACCTTGCAGGAAAAAGGGCAAAACAGGATTTTCAGCTTT 2353
 Qy 341 IleIleLysGlnAseGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 2354 ATTATTAGCCAAATGGTATGTTCTCATTCAGTGGCTGACAAAGAACAGTAGAACGT 2413
 Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAseValAlaGlyMet 380
 Db 2414 CTGCGTATGATGATGTTGGTATATACGCTGTCAGTCCGTCGTATTACGTCGCTGGCTG 2473
 Qy 381 ThrProAsePseMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 2474 ACGTAGAAAAACATGGCTCCTCATCTATGTGAAGCCATTGTTGCAGTACTC 2521
 RESULT 12
 ADT46374
 ID ADT46374 standard; cDNA; 1188 BP.
 XX
 AC ADT46374;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polynucleotide #21125.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polynucleotide; gene; ss.
 XX
 OS Bacteria.
 XX
 PN US2003233675-A1.
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 DR
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 4812; 122pp; English.
 XX

CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

SQ Sequence 1188 BP; 340 A; 268 C; 269 G; 311 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4,478-162 Length: 1188
 Score: 1654.00 Matches: 310
 Percent Similarity: 86.9% Conservative: 34
 Best Local Similarity: 78.3% Mismatches: 52
 Query Match: 80.9% Indels: 0
 DB: 13 Gaps: 0

US-10-673-786A-2 (1-396) x ADT46374 (1-1188)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 1 ATGTTGAAATAATTCAGACGACCTGCCGACCTATCTCTGTTAGCGGATAGCTTT 60
 Qy 21 ArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValTyLysAspGluThr 40
 Db 61 AAAGCTGATCTCGTGAAATAATAATCAATCTGGGTATCGGTGTTTACAAAGACGAAACT 120
 Qy 41 GlyIleThrProValLeuThrSerValIleLysAlaGluInTyLysLeuGluAenGlu 60
 Db 121 GGTAACACCCCTGCTGACCAACCGTTAAATAAGCAGAAATAATCTCTGCGAAACGAA 180
 Qy 61 ThrThrLysAenTyLysLeuGlyIleAspGlyIleProGluPheGlyValGlyCysThrGlnGlu 80
 Db 181 ACCACCAATAATATCTGGCAATTAAGTGGTACTGATTTGGCGGCTACTCAGGAA 240
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
 Db 241 CTACTCTTCGGCAATACCGACGACTATCATCTAGCAAAACGCGCGGTACTGTACAAAGC 300
 Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
 Db 301 CCAGCGGAAACCGGTGCACTGCGCACTGCGCAGATTTTATTTGCTAAACACGACCAATGCA 360
 Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
 Db 361 AAACGTGTCTGATCAGTAACCAATGCGCAATGCGCAATGCGGCTTTCTCTAGCGCT 420
 Qy 141 GlyLeuGluValArgGluTyAlaTyTrpAspAlaGluAenHisThrLeuAspPheAsp 160
 Db 421 GGGTAGAATATCCCGAGTACCAATATACAAATGCGAAGACGATGCTCTGATTTTGCAT 480
 Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 481 GGCATGTTGGTGTCTGCTGAAGCTCAAGCAGGTGATGTTGTTCTCTGCTTACGCGTTGC 540

Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 Db 541 TGCCATAACCAACCGGTATCGATCCACAGCGCTGAACAGTGGCAGAAACTTGCAGATTTA 600
 Qy 201 SerValGluTyGlyTrpLeuProLeuPheAspPheAlaTyTrpGlnGlyPheAlaArgGly 220
 Db 601 TCTCGGCAAAATGCGTGGCTGCTGTTGATTTGCTTATCAGGGTGTTCGCCAAGT 660
 Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 Db 661 TTAGATGAAGATGCGAAGGCTGCGTATTTTACTAAACCAATAAGTATGATCGTT 720
 Qy 241 AlaSerSerTyTrpSerLysAenPheGlyLeuTyTrpAenGluArgValGlyAlaCysThrLeu 260
 Db 721 GCCAGCTTCTACTCCAAATTTTGGCTGTACAAATGAGCGGTGTAGGTGCGCTGCCAATC 780
 Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 781 GTTGCAACTGACAGCGATACCGCAGAAAGCAATTCAGTCAGCGCAAAATCTATTGTTCT 840
 Qy 281 AlaAenTyTrpSerAenProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
 Db 841 ACTAACTACTCTAACCCACCGCACCGCGCTTCTGTTGTACTACAAATCTGTCCAAAT 900
 Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 901 GACGAATTTAAAGCAGAAATGATTCAGAACTGCGCACTATCGTGAACGCAATCGTCT 960
 Qy 321 MetArgGlnLeuPheValAenThrLeuGlnGlyLysGlyAlaAenArgAspPheSerPhe 340
 Db 961 ATGCGTCAAGTGTGTTGTGAACACCTTGCAAGAAAGGTGCAAAACACAGACTTTAGCTTT 1020
 Qy 341 IleIleLysGlnAenGlyMetPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1021 ATTATTAGCCAAATGATGTTTTCATCTAGTGTCTGACTAAAGAACAGTTCATCT 1080
 Qy 361 LeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAenValAlaGlyMet 380
 Db 1081 CTGCGTGAAGAAATTCGGTATTTATGCGGTACGCTCGGCGCATTAACGTAGCGGCTTA 1140
 Qy 381 ThrProAenAenMetAlaProLeuGluAlaIleValAlaValLeu 396
 Db 1141 ACATCGGAGAACTGGTTCTCTGTGTGAAGCGATTTGTCAGTACTC 1188

RESULT 13
 ADT05500
 ID ADT05500 standard; DNA; 7977 BP.
 XX
 AC ADT05500;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Haemophilus influenzae (NTHi) contig DNA sequence - SEQ ID 536.
 XX
 KW middle ear bacterial infection; nasopharynx bacterial infection; ds;
 XX
 OS Haemophilus influenzae.
 XX
 PN WO2004078949-A2.
 XX
 PD 16-SEP-2004.
 XX
 PF 05-MAR-2004; 2004WO-US007001.
 XX
 PR 06-MAR-2003; 2003US-0453134P.
 XX
 PA (CHIL-) CHILDRENS HOSPITAL INC.
 XX
 PI Bakaletz LO, Munson RS, Dyer DW;
 XX
 DR WPI; 2004-662422/64.
 XX

PT New polynucleotides of nontypeable strain of Haemophilus influenzae,
 PT useful for treating or preventing NTHi bacterial infections of the middle
 ear and/or nasopharynx.

PS Example 1; SEQ ID NO 536; 88pp; English.

CC The invention comprises nucleotide sequences (genes) from the genome of a
 CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
 CC sequences of the invention are useful for treating or preventing NTHi
 CC bacterial infections of the middle ear and/or nasopharynx. The present
 CC nucleic acid represents an NTHi contig sequence of the invention.

XX Sequence 7977 BP; 2664 A; 1592 C; 1464 G; 2255 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 1,51e-128 Length: 7977
 Score: 1343.00 Matches: 251
 Percent Similarity: 78.3% Conservative: 59
 Best Local Similarity: 63.4% Mismatches: 86
 Query Match: 65.7% Indels: 0
 DB: 13 Gaps: 0

US-10-673-786A-2 (1-396) x ADT05500 (1-7977)

Qy 1 MetPheGluAenIleThrAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
 Db 1594 ATGTTGAACATATTAAAGCGGACCGGATCCATCTTAGGCTTAGGCGAAGCATTC 1653
 Qy 21 ArgAlaaspGluAargProGlyLysIleAenLeuGlyLeuGlyValTyrLysAaspGluThr 40
 Db 1654 AAATCCGAAACTCCGAAAAATAAATTAACTTGGGTATTGGCGTTTATAAAGATGCCCAA 1713
 Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
 Db 1714 GGCACACCCCGATATGCGTGCAGTGAAGACGACGAAACCCCTATTATTGTAAGGAA 1773
 Qy 61 ThrThrLysAenTyrLeuGlyIleAaspGlyLeuProGluPheGlyA:GcYsThrGlnGlu 80
 Db 1774 AAAACGAAATATTATCTGACTATCGATGTTGCTGATTATACGAACAAACAAAGCG 1833
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaAargThrAlaGlnThr 100
 Db 1834 TTACTTTTCGGTAAAGATTCTGAAGTTCATCAATCTAATCGTGCARGAACAGTACAAAGT 1893
 Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAaspPheLeuAlaLysAenThrSerVal 120
 Db 1894 TTAGCGGACAGCGGCATACGATTCGCGCAGATTTATTAACGTCACAACTAAGCA 1953
 Qy 121 LysArgValTyrValSerAsnProSerTrpProAenHisLysSerValPheAenSerAla 140
 Db 1954 CAAAATGTTTGGATCAGCACCCCAACTTGGCCAAACCAACCAATGCTTTTCAATGCCGTC 2013
 Qy 141 GlyLeuGluValArgGluTyrAlaTyrThrAapAlaGluAenHisThrLeuAaspPheAasp 160
 Db 2014 GGTATGCAATTCGTGAATATCGTATTATGATGCTGAACGCAAAAGCCCTTTGATGGGAT 2073
 Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCys 180
 Db 2074 CATTACTTGAAGATTTAAGCAACGACGAGCGATGCGTCTTTACACGGTTGT 2133
 Qy 181 CysHisAenProThrGlyIleAaspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
 Db 2134 TGCCATAATCCGACTGGTATTGACCCCTACTCCAGACAAATGGCAAGATTAGCCGCACTT 2193
 Qy 201 SerValGluLysGlyTyrLeuProLeuPheAaspPheAlaTyrGlnLysPheAlaArgGly 220
 Db 2194 TCGCGAAAAACGGCTGGTGGCACTCTTGACTTTGCTTATCAAGTTTAGGCAATGGA 2253
 Qy 221 LeuGluGluAaspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
 Db 2254 TTAGATGAAGATGCTTATGGCTTGGGCTTTTGGCGCAACCAACCAAGAATTTGTAGTA 2313
 Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluAargValGlyAlaCysThrLeu 260

Db 2314 GCAGATTCATCTCGAAAAAACTTTGGTTATATATAAGAACTGTGGTGCCTTACTCTT 2373
 Qy 261 ValAlaAlaAaspSerGluThrValAaspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 2374 GTGCAGAAAATGCAGAAATTCATCAACCGCATTAACCCCAAGTGAATCAATATTTCGT 2433
 Qy 281 AlaAenTyrSerAenProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
 Db 2434 ACATCTTACTCTTAACCTCGATCTCACGGCGGCGCAACTGTAGCAACAGTATTAAATGAC 2493
 Qy 301 AspAlaLeuArgAlaIleTyrGluGlnLeuThrAaspMetArgGlnArgIleGlnArg 320
 Db 2494 ACTCAACTTCGCCAAGAGTGGGAAATGAATTAACGAAATGCCGCAACGATCAAAAA 2553
 Qy 321 MetArgGlnLeuPheValAenThrLeuGlnGluLysGlyAlaAenArgAaspPheSerPhe 340
 Db 2554 ATCGTCATTTATTCGTTTCAGTTATTAAAGAATACGGTCGGGCAACAGATTTTCAGCTTT 2613
 Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 2614 ATCATTGAACAAAAACGATATGTTCTCTTTCAGTGGATTAACTGGGGAACAGTGATCGC 2673
 Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenValAlaGlyMet 380
 Db 2674 TTAAGAAGAAGAAATTCGCAATTTATGCTGTTTCGTTGCTATCAATGTCGCTGGTATT 2733
 Qy 381 ThrProAaspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 2734 ACTGAAGATTAATTCGCTATTATGTGAAGCATTGTGAAGAGTCTT 2781

RESULT 14

ADT05649/c
 ID ADT05649 standard; DNA; 349980 BP.

XX AC ADT05649;

XX DT 02-DEC-2004 (first entry)

XX DE Haemophilus influenzae (NTHi) DNA sequence - SEQ ID 685.

XX DE middle ear bacterial infection; nasopharynx bacterial infection; ds.

XX OS Haemophilus influenzae.

XX PN WO2004078949-A2.

XX PD 16-SEP-2004.

XX PF 05-MAR-2004; 2004WO-US007001.

XX PR 06-MAR-2003; 2003US-0453134P.

XX PA (CHIL-) CHILDRENS HOSPITAL INC.

XX PI Bakaletz LO, Munson RS, Dyer DW;

XX DR WPI; 2004-662422/64.

XX PT New polynucleotides of nontypeable strain of Haemophilus influenzae,
 useful for treating or preventing NTHi bacterial infections of the middle
 ear and/or nasopharynx.

XX PS Claim 1; SEQ ID NO 685; 88pp; English.

XX CC The invention comprises nucleotide sequences (genes) from the genome of a
 CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
 CC sequences of the invention are useful for treating or preventing NTHi
 CC bacterial infections of the middle ear and/or nasopharynx. The present
 CC nucleic acid represents an NTHi DNA sequence of the invention.

XX SQ Sequence 349980 BP; 105127 A; 63538 C; 70035 G; 111276 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 2,92e-126 Length: 349980
 Score: 1343.00 Matches: 251
 Percent Similarity: 78.3% Conservative: 59
 Best Local Similarity: 63.4% Mismatches: 86
 Query Match: 65.7% Indels: 0
 DB: 13 Gaps: 0

US-10-673-786A-2 (1-396) x ADT05649 (1-349980)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGluAlaAspLeuPhe 20
 Db 217833 ATGTTGAACATATTAAAGCGGCACAGCCGATCCAACTTAGGCTTAGCGAAGCATTC 217774

QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThr 40
 Db 217773 AAATCCGAACCTCGGAAATTAATTAATCTGGGTATGGCTTTAAGATCGGCA 217714

QY 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyLysLeuGluAsnGlu 60
 Db 217713 GGCACACCCCGATTATGTCAGTGAAGAGCAGAAACAGCTTATTGTAAGGAA 217654

QY 61 ThrThrLysAsnTyLysLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 217653 AAAACGAAAAATTTATCTGACTATCGATGCTATTTGCTGATTATAACGACAAACAAAGCG 217594

QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysAlaArgThrAlaGlnThr 100
 Db 217593 TTACTTTTCGGTAAGATTCTGAAGTCATCAATCTAATCTGTCGCAAGAACAGTACAAAGT 217534

QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 217533 TTAGCGGGAACAGCGCATTCAGCATTCGCGCAGAAATTATTAAACGTCAAACTAAAGCA 217474

QY 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 Db 217473 CAATAATGTTGGATCAGCACCCCACTTGGCCAAACCAATGCCATTTTCAATCCGTC 217414

QY 141 GlyLeuGluValArgGluTrpAlaTyLysAspAlaGluAsnHisThrLeuLeuAspPheAsp 160
 Db 217413 GGTATGACAAATTCGGTAATATCGTTATTATGATCTGAACGCAAGCCCTTGATGGGAT 217354

QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 217353 CATTTACTTGAGATTAAAGCCAGCAAGCGAGCGATGTGCTTTTACACGGTGT 217294

QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 Db 217293 TGCCATAATCCGACTGGTATTGACCCCTACTCCAGAACCAATGCGAAGAAATTAGCCGCACTT 217234

QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyLysGlnGlyPheAlaArgGly 220
 Db 217233 TCGCGGAAAAACGGCTGGTGGCTCTTTGACTTTTGTCTTATCAAGGTTTAGCCAAATGA 217174

QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 Db 217173 TTAGATGAAGATGCTTATGGCTTGGCTTTGGCGCAACCAACCAAGAAATTTAGTA 217114

QY 241 AlaSerSerTyLysAsnPheGlyLeuTyLysAsnGluArgValAlaGlyCysThrLeu 260
 Db 217113 GCGAGTTCATCTCGAAAAACCTTGGTTTATATATGAACGTTGTTGGCTTTACTCTT 217054

QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 217053 GTGCGAGAAAAATGCGAATAATTCATCAACCGCATTAACCCCAAGTGAATATTATTCGT 216994

QY 281 AlaAsnTyLysSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 Db 216993 ACACCTCTACTTAACCTCGCATCTCAGCGCGCGCAACTGTAGCAACAGTATTAAATGAC 216934

QY 301 AspAlaLeuArgAlaIleTrpGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 216933 ACTCAACTTCGCCAAGAGTGGGAAAAATGAATTTAACTGAATTCGCCAAGCATTCAAAAA 216874

QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
 Db 216873 ATGCGTCATTATTTCAGTTATTAAAGAAATACCGTGGCGAAACAAGATTTCAGCTTT 216814

QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 216813 ATCATTTGAACAAAACGGTATGTTCTTTTCAGTGGATTAACTGGGGAACAAGTGGATCGC 216754

QY 361 LeuArgGluGluPheGlyValTyLysAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 216753 TTAAAGAGAAATTTGCCATTATTGCTTTCGTTCTGCTCGTATCAATGCTGCTGTTAT 216694

QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 216693 ACTGAAGATAATATTTCGCTATTATGTGAAGCATTTGTGAAGTGTCTT 216646

RESULT 15

AAT42063_16
 Continuation (17 of 19) of AAT42063 from base 1600001 (Haemophilus influenzae complete
 WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063

WP	Fragment Name	Begin	End
WP	AAT42063_00	1	110000
WP	AAT42063_01	100001	210000
WP	AAT42063_02	200001	310000
WP	AAT42063_03	300001	410000
WP	AAT42063_04	400001	510000
WP	AAT42063_05	500001	610000
WP	AAT42063_06	600001	710000
WP	AAT42063_07	700001	810000
WP	AAT42063_08	800001	910000
WP	AAT42063_09	900001	1010000
WP	AAT42063_10	1000001	1110000
WP	AAT42063_11	1100001	1210000
WP	AAT42063_12	1200001	1310000
WP	AAT42063_13	1300001	1410000
WP	AAT42063_14	1400001	1510000
WP	AAT42063_15	1500001	1610000
WP	AAT42063_16	1600001	1710000
WP	AAT42063_17	1700001	1810000
WP	AAT42063_18	1800001	1830121

Alignment Scores:

Pred. No.: 4.37e-125 Length: 110000
 Score: 1325.00 Matches: 248
 Percent Similarity: 77.8% Conservative: 60
 Best Local Similarity: 62.6% Mismatches: 88
 Query Match: 64.8% Indels: 0
 DB: 2 Gaps: 0

US-10-673-786A-2 (1-396) x AAT42063_16 (1-110000)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 84147 ATGTTGAACATATTAAAGCGGCACAGCCGATCCAACTTAGGCTTAGCGAAGCATTT 84206

QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThr 40
 Db 84207 AAATCCGAACCTCGGAAAAATTAATCAATTTGGGTATGGCTTTATTAAGATCGGCA 84266

QY 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyLysLeuGluAsnGlu 60
 Db 84267 GGCACACCCCAATTTATGACCGCGTAAAGAGCGGAAAAACAGTATTATTGTAAGGAA 84326

QY 61 ThrThrLysAsnTyLysLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 84327 AAAACCAAGAATTTATCTGACTATCGATGCTATTTGGGTATTTGGCTTTATTAAGATCGGCA 84386

QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysAlaArgThrAlaGlnThr 100
 Db 84387 CTCCTTTTCGGTAAAGATTCTGAAGTCATCCATCTAATTCGAGCAAGAACAGTACAAAGT 84446

QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120

Db	84447	TTAGGCGGAACAGGTGCATTACGATTGGCGAGAAATTTATTAACACGCCAACTAAAGCA	84506
Qy	121	LYsArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla	140
Db	84507	CAAAATGTTTGGWTCAGCACACCACTTGGCCAAACCAACATGCGATTTTCATGCTGTC	84566
Qy	141	GlyLeuGluValArgGluTyfAlaTyfTyfAspAlaGluAsnHisThrLeuAspPheAsp	160
Db	84567	GGTATGACCAATTCGTGAATATCGTTATATGATGCTGAACGCAAGCCCTTGATTGGGA	84626
Qy	161	AlaIleuLeaAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys	180
Db	84627	CACATTATTAGAAGATTTAAGCCCAAGCAAGCGATGTGGTGCTTTTACACGGTTGT	84686
Qy	181	CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu	200
Db	84687	TGCGCATATCCGACTGGTATTGACCCCTACTCCAGAACAAATGGCAGAAATTAGCCGCA	84746
Qy	201	SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyfGlnGlyPheAlaArgGly	220
Db	84747	TCAGCTAAAAATGGTTGGTTGCCACTCTTTGACTTTTGCTTTATCAAGGTTTAGCCAA	84806
Qy	221	LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal	240
Db	84807	TTAGATGAAGATGCTTATGGTTTACGTGCTTTTGACGCAAAACCCACAAAGAATTA	84866
Qy	241	AlaSerSerTyrSerLysAsnPheGlyLeuTyfAsnGluArgValGlyValaCysThrLeu	260
Db	84867	GCAGGTTTCATCTCGAAAAACTTGTGTTTATATATGAACGTGTGTGTGCATTACCC	84926
Qy	261	ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg	280
Db	84927	GTGCGAATAATGCAGAAATTTGCATCAACTCATTAACCAAGTAAATCAATTATTCG	84986
Qy	281	AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn	300
Db	84987	ACCTTATACTCTAACCCAGCTTCTCAGCGGGGCGACCGTAGCAACAGTATTAATGAT	85046
Qy	301	AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg	320
Db	85047	GCTCAACTTCGCCAAGATGGGAAATTAATTAATGAAATCGGTGAACGATCAAAAA	85106
Qy	321	MetArgGlnLeuPheValAsnThrLeuGlnGlnLysGlyAlaAsnArgAspPheSerPhe	340
Db	85107	ATCGCTCACATTATTCGTTTCAGTTTATTAAGAATAATGTGTGCAGAACAAAGATTC	85166
Qy	341	IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg	360
Db	85167	ATCATTTGAACAAACCGTATGTTTCTTCAGTGGATTAACAGGGGGAACAAGTGGATCG	85226
Qy	361	LeuArgGluGluPheGlyValTrpAlaValAlaSerGlyArgValAsnValaLaglyMet	380
Db	85227	TTAAAAAATGAATTTGCCATTACGCTGTTCGTTCTGTGCTGATCAAGTAGCTGGAATC	85286
Qy	381	ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu	396
Db	85287	ACAGAAGATAATATTCGCTACTTATGTGAAGTATCGTGAAAGTACTT	85334

Search completed: March 15, 2006, 03:34:59
Job time : 1478 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 15:18:01 ; Search time 5196 Seconds
(without alignments)
4332.180 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045

Sequence: 1 MFENITAAPDPILGLADLF.....VAGMTDPDNPALCEAIVAVL 396

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSSWEB/spool/US10673786/runat_14032006_135609_22591/app_query.fasta_1
-DB=genEmbl -QFMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs03p
-USER=US10673786_QCGN_1_1_7415 @runat_14032006_135609_22591 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_ro.*

10: gb_sts.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2045	100.0	1293	6	I08485 Sequence 12
2	2045	100.0	1331	1	ECASPCG X05904 E. coli asp
3	2045	100.0	1331	6	AR095441 Sequence

4	2045	100.0	1331	6	E16758	E16758 gDNA encodi
5	2045	100.0	1415	1	ECASPC	X03629 E. coli asp
6	2045	100.0	1468	6	I08484	I08484 Sequence 9
7	2045	100.0	3659	6	I08487	I08487 Sequence 15
8	2045	100.0	11470	1	D90731	D90731 Escherichia
9	2045	100.0	14759	6	AX370270	AX370270 Sequence
10	2045	100.0	20604	1	D90730	D90730 Escherichia
11	2045	100.0	110000	1	U00096	U00096 Escherichia
12	2037	99.6	110000	1	AE005174_11	AE005174_11
13	2037	99.6	110000	1	AE005174_12	AE005174_12
14	2037	99.6	110000	1	BA000007_11	BA000007_11
15	2033	99.4	301276	1	AE016758	AE016758 Escherich
16	2026	99.1	110000	1	AE005674_09	AE005674_09
17	2026	99.1	292504	1	AE016981	AE016981 Shigella
18	1982	96.9	62027	1	AE008743	AE008743 Salmonell
19	1979	96.8	110000	1	AE017220_10	AE017220_10
20	1978	96.7	110000	1	CP000026_18	CP000026_18
21	1978	96.7	294050	1	AL627268	AL627268 Salmonell
22	1978	96.7	301983	1	AE016840	AE016840 Salmonell
23	1931	94.4	1209	6	AR388173	AR388173 Sequence
24	1780	87.0	110000	1	BX950851_28	BX950851_28
25	1775	86.8	14855	1	AE013879	AE013879 Yersinia
26	1775	86.8	110000	1	EX936398_17	EX936398_17
27	1775	86.8	210050	1	AJ414148	AJ414148 Yersinia
28	1775	86.8	291847	1	AE017131	AE017131 Yersinia
29	1682	82.2	249878	6	AX770903	AX770903 Sequence
30	1682	82.2	349071	1	BX571864	BX571864 Photorhab
31	1349	66.0	9431	1	AY795600	AY795600 Actinobac
32	1343	65.7	7977	6	CQ872923	CQ872923 Sequence
33	1343	65.7	110000	1	CP000057_13	CP000057_13
34	1343	65.7	349980	6	CQ873072	CQ873072 Sequence
35	1336	65.3	110000	1	CP000020_16	CP000020_16
36	1333	65.2	348672	1	CR378670	CR378670 Photobact
37	1329	65.0	10952	1	U32835	U32835 Haemophilus
38	1325	64.8	10766	1	AE015676	AE015676 Shewanell
39	1325	64.8	110000	1	BA000031_19	BA000031_19
40	1325	64.8	110000	6	BD426631_16	BD426631_16
41	1325	64.8	110000	6	AR274513_16	AR274513_16
42	1325	64.8	110000	6	AF632719_16	AF632719_16
43	1324	64.7	7831	1	AF183931	AF183931 Aeromonas
44	1318	64.4	300029	1	AE016804	AE016804 Vibrio vu
45	1317	64.4	110000	1	AE004969_14	AE004969_14

ALIGNMENTS

RESULT 1

I08485

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .1293

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.: 7.8e-149

Score: 2045.00

Percent Similarity: 100.0%

Best Local Similarity: 100.0%

Query Match: 100.0%

DB: 6

Length: 1293

Matches: 396

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

linear PAT 02-DEC-1994

1293 bp DNA

Sequence 12 from Patent WO 8700202.

I08485

I08485

I08485.1 GI:588805

Unknown.

Unknown.

Unclassified.

1 (bases 1 to 1293)

Edwards, M.R., Taylor, P.P., Hunter, M.G. and Fotheringham, I.G.

COMPOSITE PLASMIDS FOR AMINO ACID SYNTHESIS

Patent: WO 8700202-A 12 15-JAN-1987;

Location/Qualifiers

1. .1293

/organism="unknown"

/mol_type="unassigned DNA"


```
/note="t was g in [1]"
/citation={1}
```

ORIGIN

Alignment Scores:

Pred. No.:	8,08e-149	Length:	133
Score:	2045.00	Matches:	396
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	1	Gaps:	0

US-10-673-786A-2 (1-396) x ECASPCG (1-1331)

Qy	1	MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyIleAlaAspLeuPhe	20
Db	10	ATGTTTGGAGAACATTACCGCCCTCTCCGCACCGGATTTCTGGCGCTTGGCGATCTGTTT	69
Qy	21	AtqAlaAspGluArgProGlyIleAlaLeuGlyIleGlyValTyrLysAspGluThr	40
Db	70	CGTGCCGATGAACCGTCCCGCAAAATTAACCTCGGGATTTGGTGCTATAAAGATGAAGC	129
Qy	41	GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu	60
Db	130	GGCAAAACCCCGGTACTGACACCGTGAAAGAAGCTGAAACAGTATCTGCTCGAAATGAA	189
Qy	61	ThrThrLysAsnTyrIleuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu	80
Db	190	ACCACCAAAATTAACCTCGCATTCAGCGCATCCCTGAATTTGGTCGCTGCACTCAGAA	249
Qy	81	LeuLeuPheGlyLysGlySerAlaIleuLeuAsnAspLysArgAlaArgThrAlaGlnThr	100
Db	250	CTGCTGTTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCGACGGCACAGCT	309
Qy	101	ProGlyGlyThrGlyValAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal	120
Db	310	CGGGGGGGACCTGGCGCATACCGCTGCTGCCGATTTCTGGCAAAATAATACACGGCTT	369
Qy	121	LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla	140
Db	370	AAGCGTGTGTGGTGAGCAACCAAGCTGGCGCAACCATTAAGAGCGCTTTTAACCTCGCA	429
Qy	141	GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp	160
Db	430	GGTCTGGAAAGTTTCGTGAATACGCTTATTATGATGCGGAAATACACACTCTTTGATTCGAT	489
Qy	161	AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys	180
Db	490	GCACGTATTAAACAGCCCTGAATGAAGCTCAGGCTGGCGACGTAGTGCTTTCCATGGCTGC	549
Qy	181	CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu	200
Db	550	TGCCATAACCCAAACCGGTATCGACCTTACGCTGGGCAACATGGCAAAACACTGGCAAACTC	609
Qy	201	SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly	220
Db	610	TCCGTTTGAGAAAGGCTGGTTACCGCTGTTTGACTTCGCTTACCAGGCTTTTGGCCGTGT	669
Qy	221	LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal	240
Db	670	CTGGGAAGAAGATGCTGAAGAGCTGCGCGCTTTCCGCGCTATGCATAAAGAGCTGATTTGT	729
Qy	241	AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu	260
Db	730	GCCAGTTCCTACTCTATAAAACTTTTGGCCCTGTACACGAGCGGTGTGGCGCTTGTACTCTG	789
Qy	261	ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaIleLeuArg	280
Db	790	GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCCAATGAAGCGCGCATTCGC	849
Qy	281	AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValVallalaThrIleLeuSerAsn	300
Db	850	GCTAACTACTTAAACCCACGACACACGGCGCTTCTGTTGTTGGCCACCATCTCTGAGCAAC	909

Qy	301	AspAlaLeuArgAlaIleTrrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg	320
Db	910	GATGCGGTACGTGCGATTTTGGGAACAAGAGCTGACTGATATGCCCGCAGCGTATTACAGCGT	969
Qy	321	MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe	340
Db	970	ATGCGTCAGTTGTTGTCATAGCTGTCAGAAAAAGGCGCAAAACCGCGACTTCAGCTTT	1029
Qy	341	IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg	360
Db	1030	ATCATCAAAACAGAACGCGCATGTTCTCTTCAGTGGCGTCAAAAAGAAACAAGTCTCGCT	1089
Qy	361	LeuArgGluGluPheGlyValTyrlaValaValaSerGlyArgValaAsnValaIleGlyMet	380
Db	1090	CTGCGCGAAAGAGTTTGGCGTATATGCGGTTGCTTCCTGGTCCGCTAAATGTGGCGGGATG	1149
Qy	381	ThrProAspAsnMetAlaProLeuCyGluAlaIleValaIleValaValLeu 396	
Db	1150	ACACCATTAACATGGCTCCGCTGTGGAAGCGGATTGTGGCAGTGTCTG 1197	
RESULT 3			
AR095441			
LOCUS			
DEFINITION Sequence 23 from patent US 6004773.			
ACCESSION AR095441			
VERSION AR095441.1 GI:10023318			
KEYWORDS			
SOURCE Unknown.			
ORGANISM Unknown.			
REFERENCE			
1 (bases 1 to 1331)			
AUTHORS Araki,M., Sugimoto,M., Yoshihara,Y. and Nakamatsu,T.			
TITLE Method for producing L-lysine			
JOURNAL Patent: US 6004773-A 23 21-DEC-1993;			
FEATURES			
source			
1. .1331			
/organism="unknown"			
/mol_type="unassigned DNA"			
ORIGIN			
Alignment Scores:			
Pred. No.:	8.08e-149	Length:	1331
Score:	2045.00	Matches:	396
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0
US-10-673-786A-2 (1-396) x AR095441 (1-1331)			
Qy	1	MetPheGluAsnIleThrAlaIleProAlaAspProIleLeuGlyLeuAlaAspLeuPhe	20
Db	10	ATGTTTGGAGAACAATTACCGCGCTCTCTGCCGACCCGATTCCTGGGCCCTGGCCGATCTGTTTT	69
Qy	21	ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrlsAspGluThr	40
Db	70	CGTGCCGATGAACTCCCGCAAAATTAACCTCGGATTGGTGTCTATAAGATGAGACG	129
Qy	41	GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrlLeuLeuGluAsnGlu	60
Db	130	GGCAAAACCCCGGTACTGCCAGCGTGAAAAAGCGTGAAACAGTATCTGCTCGCAAAATGAA	189
Qy	61	ThrThryAsnTyrlLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu	80
Db	190	ACCACCAAAATTAACCTCGGCATTTGACCGCATCCCTGGAATTTGGTCCGCTGCATCAGGAA	249
Qy	81	LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr	100
Db	250	CTGCTGTTTGGTAAAGTAGGCGCCCTGATCAATGACAAACGTGCTCGCAGCGCACAGACT	309
Qy	101	ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal	120

Db 310 CCGGGGGGCACTGGCGCCTACTACGCTGGCTGGCGATTTCTCGCAAAAATACCAGCGTT 369

Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140

Db 370 AAGCGTGTGGTGAGCAACCAAGCTGGCGCAACCAATAAGAGCGCTTTAACTCTGCA 429

Qy 141 GlyLeuGluValArgGluTrpAlaTyTrpAenHisLysSerValPheAenSerAla 160

Db 430 GGTCTGGAAGTTCGTGAATACGCTTATTATGATGGGAAATACACACTCTTACCTTCGAT 489

Qy 161 AlaLeuLeuAenSerLeuAenGluAlaGlnAlaGlyAenValValLeuPheHisGlyCys 180

Db 490 GCACCTGATTAACAGCTGATGAGCTAGGCTAGGCTGGGACGTAGTGTCTTCCATGGCTGC 549

Qy 181 CysHisAenProThrGlyLeuAenProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200

Db 550 TGCCATAACCAACCGGTATCGACCTACGCTGGAACAATGGCAACACTGGCACAACCTC 609

Qy 201 SerValGluGlyTrpLeuProLeuPheAenPheAlaTyGlnGlyPheAlaArgGly 220

Db 610 TCCGTTGAGAAAGGCTGGTTACCGCTGTTTGACTTCGCTTACAGGGTTTGGCCGTGGT 669

Qy 221 LeuGluGluAenAlaGluGlyLeuArgAlaPheAlaAenMetHisLysGluLeuVal 240

Db 670 CTGGAGAGAGTGTGAAGGACTGGCGGCTTTCGGGCTATGCATTAAGAGCTGATTGTT 729

Qy 241 AlaSerSerTrpSerLysAenPheGlyLeuTrpAenGluArgValGlyAlaCysThrLeu 260

Db 730 GCCAGTTCCTACTCTAAAACCTTGGCCCTGTACACGAGCGTGTGGCGCTTGTACTCTG 789

Qy 261 ValAlaAenSerGluThrValAenArgAlaPheSerGlnMetLysAlaAlaArg 280

Db 790 GTTGCTGCCGACAGTGAAACCGTTGATCGCGCATTCAGCCAAATGAACGGCGATTCGC 849

Qy 281 AlaAenTrpSerAenProAlaHisGlyAlaSerValValAlaThrileLeuSerAen 300

Db 850 GCTAACTACTCTAACCCACAGCACACGGCGCTTCTGTTGGCCACCATCTGAGCAAC 909

Qy 301 AspAlaLeuArgAlaAlaTrpGluGlnGluLeuThrAspMetArgGlnArgileGlnArg 320

Db 910 GATCGCTTACGTGGCGATTTGGGAACAAGAGCTGACTGATATGCGCAGCGTATTCAGCGT 969

Qy 321 MetArgGlnLeuPheValAenThrLeuGlnLysGlyAlaAenArgAspPheSerPhe 340

Db 970 ATGCGTCAGTTGTCTGCTCAATACGCTGCAGGAAAGGCGCAACCGCGACTTCAGCTTT 1029

Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360

Db 1030 ATCATCAACAGAACGGCGATTTCTCTTCAGTGGCTGACCAAGAACAGAGTGTGCGT 1089

Qy 361 LeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAenValAlaGlyMet 380

Db 1090 CTGGCGAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTGGCGGGATG 1149

Qy 381 ThrProAenAenMetAlaProLeuGlyGluAlaIleValAlaValLeu 396

Db 1150 ACACCAAGTAACATGGCTCGCTGTGCGAAGCGGATTTGTGGCAGTGCTG 1197

RESULT 4

E16758 LOCUS E16758 1331 bp DNA linear PAT 28-JUL-1999

DEFINITION gDNA encoding aspartate aminotransferase (AAT).

ACCESSION E16758

VERSION E16758.1 GI:5711441

KEYWORDS JP 1998215883-A/16.

SOURCE Escherichia coli

ORGANISM Escherichia coli

REFERENCE 1 (bases 1 to 1331)

AUTHORS Araki, M., Sugimoto, M., Yoshihara, Y. and Nakamatsu, W.

TITLE PRODUCTION OF L-LYSINE

JOURNAL Patent: JP 1998215883-A 16 18-AUG-1998;

COMMENT

AJINOMOTO CO INC

OS Escherichia coli

PN JP 1998215883-A/16

PD 18-AUG-1998

PF 03-DEC-1997 JP 1997333238

PR 05-DEC-1996 JP 96P 325659

PI ARAKI MASAYUKI, SUGIMOTO MASAKAZU, YOSHIHARA YASUHIKO, PI

NAKAMATSU WATARU

PC C12N15/09, C12N1/21, C12P13/08, (C12N15/09, C12R1:15), (C12N1/21, C12R1:15),

PC (C12P13/08, C12R1:15);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

CC Location/Qualifiers

FT Key

FT source 1..1331

FT /organism='Escherichia coli'

FT /strain='JM109'

FT 5'UTR 1..9

FT CDS 10..1197

FT /product='aspartate aminotransferase' FT

FT /gene='aspC'

FT 3'UTR 1198..1331.

FEATURES

source

1..1331

/organism='Escherichia coli'

/mol_type='genomic DNA'

/db_xref='taxon:562'

ORIGIN

Alignment Scores:

Pred. No.: 8,086-149 Length: 1331

Score: 2045.00 Matches: 396

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 6 Gaps: 0

US-10-673-786A-2 (1-396) x E16758 (1-1331)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAenProIleLeuGlyLeuAlaAenLeuPhe 20

Db 10 ATGTTTGAGAAACATTACCGCCGCTCTCGCGACCGATTCGGGCTCGCCGATCTGTTT 69

Qy 21 ArgAlaAenGluArgProGlyLysIleAenLeuGlyIleGlyValTrpLysAenGluThr 40

Db 70 CGTGGCGATGAACGTCCTCGGCAAAATTAACCTCGGGATTGGTGTCTATAAGATGAGACG 129

Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTrpLysLeuAenGlu 60

Db 130 GGCAAAACCCCGCTACTGACCAAGCGTGAAAGGCTGAACAGTATCTGCTCGAAATGAA 189

Qy 61 ThrThrLysAenTyLeuGlyIleAenGlyIleProGluPheGlyArgCysThrGlnGlu 80

Db 190 ACCACCAAAATTAACCTCGGCATTCAGCGCATCCCTGATATTTGGTTCGCTCAGCAAGAA 249

Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAenLysArgAlaAenArgThrAlaGlnThr 100

Db 250 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCCTCGCACGCGACAGACT 309

Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAenPheLeuAlaLysAenThrSerVal 120

Db 310 CCGGGGGGCACTGGCGCCTACGCGCTGGCTGGCGATTTCTCGCAAAAAATACCAGCGTT 369

Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140

Db 370 AAGCGTGTGGTGGTGGAGCAACCCAGCTGGCGCAACCATTAAGAGCGCTTTAACTCTGCA 429

Qy 141 GlyLeuGluValArgGluTrpAlaTyTrpAenHisLysSerValPheAenSerAla 160

Db 430 GGTCTGGAAGTTCGTGAATACGCTTATTATGATGGGAAATACACACTCTTACTTTCGAT 489


```

Db      618  GCACCTGATTAAACAGCTGAATGAAGCTCAGGCTGGCGAGTGTGTCTCCATGCTGC 677
Qy      181  CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db      678  TGCATTAACCCACCGGTATCGACCTAGCTGGAAACATGGCAACACTGGCACACTC 737
Qy      201  SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db      738  TCCGTTGAGAAAGGCTGGTTACCGCTGTTGACTTCGCTTACCAAGGTTTGGCCGCTGT 797
Qy      221  LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db      798  CTGGAAGAAGATGCTGAAGAGCTGGCGCTTTCGCGGCTATGCATAAAGAGCTGATTGT 857
Qy      241  AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db      858  GCCAGTTCTCTCTTAANAACCTTGGCCCTGTACACGAGCGTGTGGCGCTTGTACTCTG 917
Qy      261  ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaIleArg 280
Db      918  GTTGTCTGCCAGACAGTGAACCGTTGATCGGCATTTCAGCAAAATGAAGCGCGATTTCG 977
Qy      281  AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db      978  GCTAACTACTCTAAACCCACGAGCACACGCGCTTCTGTGTTGCCACCATCTCGAGCAAC 1037
Qy      301  AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db      1038  GATGCGTTACGTGCGAATTTGGGAACAAGAGCTGACTGATATGGCCGAGGCTATTTCAGC 1097
Qy      321  MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db      1098  ATGCGTCAGTTGTTCTGTAATACGCTGCAGGAAAAAGGCGCAAAACCGCGACTTCAGCT 1157
Qy      341  IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db      1158  ATCATCAAAACAGAAACGCGATGTTCTCTTCAGTGGCGCTGACAAAAGAACAGTGTGCG 1217
Qy      361  LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db      1218  CTGCGCGAAGAGTTTGGCGTATATGCGTTGCTTCTGTCGCGTAATGTGGCCGGGATG 1277
Qy      381  ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db      1278  ACACCAAGATAACATGCTCGCTGTCGGAAGCGATTGTGGCAGTGTGCTG 1325

```

```

RESULT 6
LOCUS      108484
DEFINITION Sequence 9 from Patent WO 8700202.
ACCESSION 108484
VERSION    108484.1
KEYWORDS   GI:588804
SOURCE     Unknown.
ORGANISM   Unknown.
REFERENCE  1 (bases 1 to 1468)
AUTHORS   Edwards,M.R., Taylor,P.P., Hunter,M.G. and Fotheringham,I.G.
TITLE     COMPOSITE PLASMIDS FOR AMINO ACID SYNTHESIS
JOURNAL   Patent: WO 8700202-A 9 15-JAN-1987;
FEATURES   Location/Qualifiers
            source
            1..1468
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.:      9,07e-149      Length:      1468
Score:          2045.00        Matches:      396
Percent Similarity: 100.0%    Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match:    100.0%      Indels:      0

```

```

DB:      6      Gaps:      0
US-10-673-786A-2 (1-396) x 108484 (1-1468)

Qy      1  MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db      191  ATGTTTGAACAATTAACCGCGCTCTCGCGACCCGATTCGGGCTTGGCCGATCTGTTT 250
Qy      21  ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db      251  CGTCCGATGAACGTCGCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 310
Qy      41  GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db      311  GGCAAAACCCCGGTACTACCAAGCGTGAAGAGGCTGAACAGTATCTGCTCGAAAATGAA 370
Qy      61  ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db      371  ACCACAAAATTAACCTCGGCATTCGCGCATCCCTGAAATTTGGTCGTCGACTCAGGAA 430
Qy      81  LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db      431  CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAAACGTCCTCGCACGCACAGCT 490
Qy      101  ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db      491  CCGGGGGGCACCTGGCGCCTACGCGTGGCTGCCGATTTCTGGCAAAAATAACACGCT 550
Qy      121  LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db      551  AAGCGTGTGGTGAGCAACCCAGCTGGCGAACCATAGAGCGCTCTTAACTCTGCA 610
Qy      141  GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db      611  GGTCTGGAAGTTCTGTTGATACGCTTATGATGCGGAAATTCACACTCTTGCATTCGAT 670
Qy      161  AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db      671  GCACGTGATTAAACAGCTGAATCAAGCTCAGCGCTGGCGAGCTAGTGTCTTCCATGGCTG 730
Qy      181  CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db      731  TGCCATAACCCACACCGGTATCGACCTAGCTGGAAACAATGGCAAAACACTGGCACACTC 790
Qy      201  SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db      791  TCCGTTGAGAAAGGCTGGTTACCGCTTTCGCTTACCGGTTTCCGCGTGTTCGCCGTGT 850
Qy      221  LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db      851  CTGGAAGAAGATGCTGAAGAGCTGGCGCTTTCGCGGCTATGCATAAAGAGCTGATTGT 910
Qy      241  AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db      911  GCCAGTTCTCTCTTAANAACCTTGGCCCTGTACACAGAGCGTGTGGCGCTTGTACTCTG 970
Qy      261  ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaIleArg 280
Db      971  GTTGTCTCGACAGTGAACCGTTGATCGCGCATTCAGCAAAATGAAGCGCGATTTCG 1030
Qy      281  AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db      1031  GCTAACTACTCTAAACCCACGAGCACACGCGCTTCTGTTGTTGCCACCATCTCGAGCAAC 1090
Qy      301  AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db      1091  GATGCGTTACGTGCGAATTTGGGAACAAGAGCTGACTGATATCGCGAGGCTATTTCAGCT 1150
Qy      321  MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db      1151  ATGCGTCAGTTGTTCTGTAATACGCTGCAGGAAAAAGGCGCAAAACCGCGACTTCAGCT 1210
Qy      341  IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360

```

Db 1211 ATCATCAACAGAACGGCATGTTCTCTTCAGTGGCTGACAAAGAAACAAAGTGGCGT 1270
 QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1271 CTGGCGGAAGAGTTTGGCGTATATGCGGTGCTTCTGTCGCGTAATGTGGCGGGAATG 1330
 QY 381 ThrProAspAsnMetAlaProLeuGlyCysGluAlaIleValAlaValLeu 396
 Db 1331 ACACCAAGATAACATGCTCGGCTGTGCGAAGCGATTGTGGCAGTGTG 1378

RESULT 7

108487
 LOCUS 108487 3659 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 15 from Patent WO 8700202.

108487
 ACCESSION 108487.1 GI:588807

108487.1
 VERSION 108487.1

Unknown.
 KEYWORDS

Unknown.
 SOURCE

Unknown.
 ORGANISM

Unclassified.
 REFERENCE

1 (bases 1 to 3659)
 AUTHORS

Edwards, M.R., Taylor, P.P., Hunter, M.G. and Fotheringham, I.G.
 TITLE

COMPOSITE PLASMIDS FOR AMINO ACID SYNTHESIS
 JOURNAL

Patent: WO 8700202-A 15 15-JAN-1987;
 FEATURES

Location/Qualifiers
 1..3659

/organism="unknown"
 /mol_type="unassigned DNA"

source

ORIGIN

Alignment Scores:
 Pred. No.: 2,68e-148 Length: 3659
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: Gaps: 6

US-10-673-786A-2 (1-396) x 108487 (1-3659)

QY 1 MetPheGluAsnIleThrAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 2344 ATGTTTGAACAATTACCGCGCTCTGCGCAGCCCGATTCTGGCCCTGGCCGATCTGTTT 2403
 QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyLeGlyValTyrLysAspGluThr 40
 Db 2404 CGTGGCGATGAACGTCCTCCGGCAAAATTAACCTCGGGATTGGTGTCTATAAAGATGAGCG 2463
 QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
 Db 2464 GGCNAACCCGGTACTGACCGCTGAAAGAGCTGAACAGTATCTGCTCGAANAATGAA 2523
 QY 61 ThrThrLysAsnTyrLeuGlyLysAspGlyLeProGluPheGlyArgCysThrGlnGlu 80
 Db 2524 ACCACCAAAAATTACCTCGGCAATTGACCGCATCCCTGAATTTGGTGGCTGCACTCAGGAA 2583
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 2584 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCGACGCGCACAGACT 2643
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 2644 CCGGGGGGCGACTGGCGCACTACCGGTGGCTGGCGATTCTTCTGGCAAAAATACACCGGT 2703
 QY 121 LysArgValTyrValSerAsnProSerTyrProAsnHisLysSerValPheAsnSerAla 140
 Db 2704 AGCGGTGTGGGTGAGCAACCAACGCTGGCGCAACCAATAGAGCGTCTTAACCTGCA 2763
 QY 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 Db 2764 GGTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAAAATCACACTCTTCACTTCGAT 2823

QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 2824 GCAGTATTAAACAGCGCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTCATGGCTGC 2883
 QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
 Db 2884 TCCATAACCCCAACCGTATCGACCTACGCTGGAAACAATGGCAACACTGGGCACAATC 2943
 QY 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 2944 TCCGTTGAGAAAGGCTGTTACCGCTGTTTGACTTACCGAGGTTTGGCCCGTGGT 3003
 QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 Db 3004 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGGCTATGCATAAAGAGCTGATGTT 3063
 QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 Db 3064 GCCAGTTCTTACTCTTAAACACTTTGGCCTGTACACGAGCGTGTGGCGCTTGACTCTG 3123
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaIleArg 280
 Db 3124 GTTGTGCGCAGCAGTGAACCGTGTGATCGCGCATTCAGCCAAATGAAGCGGATTCGC 3183
 QY 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValAlaThrLeuSerAsn 300
 Db 3184 GCTAACTACTCTAAACCCACAGCAGCGCGCTTCTGTGTGTGCGCACTCTCTGAGCAAC 3243
 QY 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 3244 GATCGTTACGTCGCGATTGCGGAACAAGAGCTGATGATGCGCCAGCGTATTACGCGT 3303
 QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyValAlaAsnArgAspPheSerPhe 340
 Db 3304 ATGCGTCAGTTGTTGCTCAATACGCTGCGAGAAAGGCGCAACCGCAGCTTCAGCTTT 3363
 QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 3364 ATCATCAACAGACAGCGCATGTTCTCTTCACTGGCGCTGACAAAGAACAAAGTGTGCGT 3423
 QY 361 LeuArgGluGlyPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 3424 CTGCGCGAAGAGTTTGGCGTATATGCGGTGCTTCTGTCGCGTAAATGTGGCCGGGATG 3483
 QY 381 ThrProAspAsnMetAlaProLeuGlyCysGluAlaIleValAlaValLeu 396
 Db 3484 ACACCAAGATAACATGCTCGCTCGCTGCGAAGCGATTGTGGCAGTGTG 3531

RESULT 8

D90731/c
 LOCUS 11470 bp DNA linear BCT 25-DEC-2002
 DEFINITION Escherichia coli K12 genomic DNA. (21.2 - 21.5 min).
 D90731 AB001340
 ACCESSION D90731.1 GI:1651452
 VERSION Complete and shotgun sequencing; ycbB; aspC; ompF; tolF; cmlB; coa;
 KEYWORDS cry; asnS; tss; pncB; pepN; ycbE; yzeB.

SOURCE

Escherichia coli K12

Escherichia coli K12

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Escherichia.

REFERENCE

1

AUTHORS

Oshima, T., Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A.,
 Ikenoto, K., Inada, T., Itoh, T., Kajihara, M., Kanai, K., Kashimoto, K.,
 Kimura, S., Kitagawa, M., Makino, K., Masuda, S., Miki, T.,
 Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishimoto, H.,
 Nishio, Y., Saito, N., Sampa, G., Seki, Y., Tagami, H., Takemoto, K.,
 Wada, C., Yamamoto, Y., Yano, M. and Horiuchi, T.

TITLE

A 718-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 12.7-28.0 min region on the linkage map

JOURNAL

PUBMED

8905232

REFERENCE

2

Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiuchi, T.,

Ikemoto, K., Inada, T., Isono, K., Isono, S., Itoh, T., Kanai, K., Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M., Kitakawa, M., Makimura, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishio, Y., Oshima, T., Saito, N., Sampei, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C., Yamamoto, Y., and Yano, M.

The systematic sequencing of the *Escherichia coli* genome in Japan Unpublished
3 (bases 1 to 11470)

Direct Submission

Mori, H.

Submitted (29-JUL-1996) Hirotsada Mori, NARA Institute of Science and Technology, Res. & Edu. Center for Genetic Info.; 8916-5

Takayama, Ikoma, Nara 630-01, Japan

(E-mail: hmori@gtc.aist-nara.ac.jp, Tel: 81-7437-2-5660,

Fax: 81-7437-2-5669)

Collaboration Information:

Project:

The Japan E. coli genome DNA sequencing project

Group:

The Japan E. coli genome DNA sequencing group

Members: (1995.4 - 1996.3)

Alba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A.,

Horiuchi, T., Ikemoto, K., Inada, T., Isono, K., Isono, S.,

Itoh, T., Kanai, K., Kasai, H., Kashimoto, K., Kim, S.,

Kimura, S., Kitagawa, M., Kitakawa, M., Makino, K.,

Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K.,

Nakamura, Y., Nishimoto, H., Nishio, Y., Oshima, T., Saito, N.,

Sampei, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C.,

Yamamoto, Y., and Yano, M.

Headed by:

Name: Takashi Horiuchi

Address: National Institute of Basic Biology, Okazaki, 444, Japan

E-mail: kishorienibb.ac.jp

Information operator:

Name: Hirotsada Mori

Address: NARA Institute of Science and Technology,

Ikoma, 630-01, Japan

E-mail: hmori@gtc.aist-nara.ac.jp

URL:

The Japan E. coli genome database

http://bsw3.aist-nara.ac.jp.

Location/Qualifiers

FEATURES source

```
1..11470
/organism="Escherichia coli K12"
/mol_type="Genomic DNA"
/strain="K12"
/db_xref="taxon:83333"
/map="21.2-21.5 min"
/clone="Kohara clone #219"
/notes="Nucleotide position 983395-994864 from the
initiation site of thra (0 min.).
This clone is from Kohara lambda miniset library."
1..47
/gene="ycbB"
```

gene

CDS

```
1..47
/gene="ycbB"
/notes="ORF_ID:0218#11
similar to PIR Accession Number PH0205"
/codon_start=3
/transl_table=11
/product="Hypothetical protein 63 (MukB 3' region)"
/protein_id="BAA35677.1"
/db_xref="GI:4062495"
/translation="SSSQIVSKAEQLIR"
228..776
```

CDS

```
/notes="ORF_ID:0219#1
similar to PIR Accession Number F64039"
/codon_start=1
/transl_table=11
/product="Hypothetical protein H11666"
/protein_id="BAA35678.1"
/db_xref="GI:4062496"
```

```
/translation="MDKFDANRRKLLALGGVALGAAILPTPAFATLSTPRPRLTLNN
LHTGESKAEFFDGRGTQBELAKLNHFRDYRANKIKSIDPGLFDQIYRLQGLGTR
KPVLLIGYSRISIDTNLNRKRSVGVAKKSYHTKGQAMDHFHIEGIALSNIRKKAALSMRA
GGVGYPRSNFVHIDTGPARHW"
```

CDS

```
803..1450
/notes="ORF_ID:0219#2
similar to SwissProt Accession Number P54501"
/codon_start=1
/transl_table=11
/product="Hypothetical 23.2 kd protein in soda-comGA
intergenic region."
/protein_id="BAA35679.1"
/db_xref="GI:4062497"
/translation="MNYRIIPVTAFSPQNSCLIWCEQTRLAALVDPGGDAEKIKQEVDD
SGLTLMQILLLTHGLHDVGAALAQHYGVFVPEKEDEFMLQGLPAQSRMFLGEEC
QPLTPDRWLNEGDTISICNTVLQVLCHPGTPGHVVFDDRAKLLISGDI PKGGVGR
SDPRGDHNLIISSIKKLLPLGDDVIFPGHGLSTLGYERLHNPFQDEMPVW"
```

gene

CDS

```
/gene="aspC"
/complement(1672..2862)
/gene="aspC"
/notes="ORF_ID:0219#3
similar to PIR Accession Number A00598"
/codon_start=1
/transl_table=11
/product="Aspartate transaminase (EC 2.6.1.1)"
/protein_id="BAA35680.1"
/db_xref="GI:1651453"
/translation="MFSNITAAPADPILGLADLFRADPRGKINLIGVYKDETKTP
VLTSVKAEQYLLENETKYNLVGIDGPEFCRQELLFGKGSALINDKRACTATPG
GTGALRVAADF LAKNTSVKRVNSNPNSHKS VFNAGLEVEYAYDAENHTLDFD
ALINSLAEQAQGDVVLHFHCCHNPTGIDPTLEQWOTLAQSVKGMFLPFDFAQGFA
RLGEDAEGLAFAMKHELIVASSYKFNLYNERVGACTLVADSETVDFRAFQOMK
AAIRANYSNPPAHGASVVATILSNLDALRAIWEQELTDMRQIRMQRLFVNTLQSKGA
NRDFSFIKQNGMFSFSLTKEQVLRLEEFVGVAVASGRVNVAGTMTFDNMAPLCEAI
VAVL"
```

gene

CDS

```
/complement(3047..4135)
/gene="cm1B"
/notes="synonyms: coa, cry, ompF, to1F"
/complement(3047..4135)
/gene="cm1B"
/notes="ORF_ID:0219#4
similar to SwissProt Accession Number P02931"
/codon_start=1
/transl_table=11
/product="Outer membrane protein f precursor (outer
membrane protein la, la, or b)."
/protein_id="BAA35681.1"
/db_xref="GI:1651454"
/translation="MKRNILAVIPALLVAGTANAAEIKYKGNKVDLYGKAVGLHY
PSKNGNSYGGNGDMTYARLGFKGETQINSDLTCYGQWYFNQNNSEGADAOQGNK
TRLAPAGIKYADVGSFDYGRNYGVYDALGYTDLMPFGGDTATSDDFVGVGVGTAT
YRNSNFGVLVDGLNFAVQYLGKNERDTRRSNGDVGGSISYEYEGFVGVAGAADR
TNLQAEPLGNGKAEQWATGKYDANNIYLAANYGETRNATPTNKNTNSGFAKNT
QDVLLVAQYQDFGLRPSIAIYTKSKADVEGIGVDLVNYPEVGATYTFNKNWSTVVD
YIINQIDSDNKLGVGDDTVAVGIYQF"
```

gene

CDS

```
/gene="aasS"
/notes="synonym: tss"
/complement(4738..6138)
/gene="aasS"
/notes="ORF_ID:0219#5
similar to SwissProt Accession Number P17242"
/codon_start=1
/transl_table=11
/product="Asparaginyl-tRNA synthetase (EC 6.1.1.22)
(asparagine-tRNA ligase) (aasRS)."
/protein_id="BAA35682.1"
/db_xref="GI:1651455"
/translation="MSVVPVADVQLQGRVADVSEVTVRGVRTRRDSKAGISPLAVYDG
SCFPBQVAVINNSLPNKNEDVLRLTTCGSLVTGKVVAS PCGQGFQFEIQASKVEVAGW
VEDPDTYPMAAKRHSIEYLRVAHLRPRTNLIGAVARVRHTLAQALHRRFNRQGFVW
STPLITASDTEGAGEMFRVSRFTLDLENLPNDQGRVDFDKDFPGKESFLTVSQNLNET
```


YACALSKIYTFGTFRFNTSRHIAEFWMLEPEVAFANLNDIAGLAEMLYKVFKA
VLEERADDMKFAERVDKDAVRLERFIEADFAQVDYTDVTLLENCKRFPENPVWG
VLSSEHYLAEEHFKAQVWVKNYPKDIKAFYMLNEDGKTVAAMDVLAPGIGELIG
GQREERLDVDERMLEMLNKEDYWWYRLRRYGVTPHSGFGLGFERLIAVVTGVQN
VRDVLPPPTBNASF"

complement (6307..7509)

/gene="pncB"

complement (6307..7509)

/gene="pncB"

/note="ORF ID:0220#1

similar to_PIR Accession Number JQ0756"

/codon_start=1

/transl_table=11

/product="Nicotinate phosphoribosyltransferase (EC

2.4.2.11)"

/protein_id="BAA35683.1"

/db_xref="GI:1651456"

/translation="MTQFASPVHLSLTDAYKLHMQQAVFHHYVDVHVAABFRCRGD
DLGIYADAIRESQVQAMHLRLODDEYQWLSALPFKADYLNWLRFRNFPQVTVSN
DNGKLDIRLSQWREVILWEVPLLAIVISEMVRYSPOQADVAQLDTSLEKLVDTSAL
TAGLDMRSRLHLDGFRFRSREVOETIVKRLQOSWFGVTSNYDLARLSLTPMGTO
AHEWFOHQOISPDLSRSORAAALAEVPPDLGIALTDCITMDAFRDFGVEFASR
YQGLRHSDSDPWESEKAIHYEKLIGIDPQSKTLVPSNLDLRKAVELYRHFSSRVQL
SFQIGRLTCDIPQVKPLNIVIKLVECNKPKVAKLSDSPGKTIChDKKAFVRLRKAFD
LPHIKKAS"

gene 7775..10387

/gene="pncB"

7775..10387

/gene="pncB"

/note="ORF ID:0220#2

similar to SwissProt Accession Number P04825"

Alignment Scores:

Pred. No.: 1,046-147 Length: 11470
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-10-673-786A-2 (1-396) x D90731 (1-11470)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaaspLeuPhe 20
DB 2862 ATGTTTGAGAAATACCGCCGCTCTGCGAGCCCGATTCTGGGCTGCGCCATCTGTTT 2803
QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrIlyAspGluThr 40
DB 2802 CGTGGCGATGAACGTCGCCGCAAAATTAACCTCGGATTGGTGTCTATAAGATGAGACG 2743
QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrIleLeuGluAsnGlu 60
DB 2742 GGCAAAACCCGGTACTACACGCGTGAAAAAGGCTGAACAGTATCTGCTCGAAAAATGAA 2683
QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
DB 2682 ACCACCAAAATATACCTGGCAATGACGCGATCCCTGAATTTGGTGGCTGCACTAGGAA 2623
QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
DB 2622 CTGCTGTTTGGTAAAGTAGCGCTGATCAATGACAAACGCTGTCGACGCGCACAGACT 2563
QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
DB 2562 CCGGGGGGCACTGGCGCACTAGCGGTGGCTGCGGATTTCTGCGAAAAAATACCAAGCGTT 2503
QY 121 LysArgValTyrValSerAsnProSerTyrProAsnHisLysSerValPheAsnSerAla 140
DB 2502 AAGCGTGTGGGTGAGCAACCAACAGCTGGCGCAACCAATGAGCGCTCTTTAACTCTGCA 2443
QY 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
DB 2442 GGTCTGGAAGTTCGTGAATACGCTTATATATGATGCGGAAAAATCACACTCTTCACTTCGAT 2383

QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
DB 2382 GCACGTGATTAA CAGCGCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTCCATGCTGC 2323
QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrPheGlnThrLeuAlaGlnLeu 200
DB 2322 TGCCATAACCCAAACCGGTATCGACCTACGCTAGCAACAATGGCAAACTCTGGCACAATC 2263
QY 201 SerValGluLysGlyTyrLeuPheAspPheAlaTyrGlnGlyPheAlaAspGly 220
DB 2262 TCCGTTGAGAAAGCTGGTTACCGCTTGTTGACTTCCTCCATTACAGAGGTTTGGCCGTGGT 2203
QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
DB 2202 CTGGAAGAAGATGCTGAAGGACTCGCGCTTTTCGGCGCTATGCATAAAGAGCTGATTGTT 2143
QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
DB 2142 GCCAGTTCCTACTCTAAAAAATTTGGCCTGTACACGAGCGGTGTGGCGCTTGTACTCTG 2083
QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
DB 2082 GTTGCTGGCGACAGTGAACCGCTTGATCGCGCATTCAGCTATGATGCCCAATGAAGCGGATTCGC 2023
QY 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValAlaThrIleLeuSerAsn 300
DB 2022 GCTAACTACTCTAACCCACGACACACGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 1963
QY 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
DB 1962 GATCGTTACGTGCGATTGGGAAACAAGAGCTGACTGATATGCCCAAGCGGATTCAGCGT 1903
QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGlyLysGlyAlaAsnArgAspPheSerPhe 340
DB 1902 ATGCGTCACTGTTGCTCAATACGCTGACGAGGAAAGCGCAACACCGACATTCAGCTTT 1843
QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
DB 1842 ATCATCAACACAGACGCGATGTTCTCTTCAGTGGCTGACAAAGAACAAGTGTGCGT 1783
QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
DB 1782 CTGCGCAAGAGCTTTGGGTATATCGGTTGCTTCTGTCGCGTAAATGTGCGCGGATG 1723
QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
DB 1722 ACACACAGATAACATGCTCGCTGCGAAGCGATTGTGCGCAGTGTGCTG 1675

RESULT 9

AX370270/c

LOCUS

Sequence 87 from Patent WO0170776.

DEFINITION

AX370270

ACCESSION

AX370270.1 GI:18857433

VERSION

KEYWORDS

SOURCE

ORGANISM

Escherichia coli

Escherichia coli

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Escherichia.

REFERENCE

1

Levy, S.B., Barbosa, T.M. and Alekshun, M.N.

Nimr compositions and their methods of use

Patent: WO 0170776-A 87 27-SEP-2001;

TRUSTEES OF TUFTS COLLEGE (US)

Location/Qualifiers

1..14759

/organism="Escherichia coli"

/mol_type="unassigned DNA"

/db_xref="taxon:562"

ORIGIN

Alignment Scores:

Pred. No.: 1.41e-147 Length: 14759
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 Db: 6 Gaps: 0

US-10-673-786A-2 (1-396) x AX370270 (1-14759)

QY 1 MetPheGluAAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 DB 2755 ATGTTTGAGAACATTACCGCCGCTCTGCCGACCCGATTCGGGCTCGCCGATCTGTTT 2696
 QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 DB 2695 CGTGCAGATGAACGTCCTCCGGCAAAATTAACTCCGGATTGGTGTCTATAAAGATGAGACG 2636
 QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
 DB 2635 GGCAGAAACCCCGGTACTGACCAAGCGTGAAAAGGCTGAACAGTATCTGCTCGAAAATGAA 2576
 QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 DB 2575 ACCACCAAAATTAACCTCGCATTCGCGCATCCCTGAATTTGGTGGCTGCACACTCAGGAA 2516
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 DB 2515 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGTGCTCGCAGGCACAGACT 2456
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 DB 2455 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTGGCAAAAATACACAGCGTT 2396
 QY 121 LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 DB 2395 AAGCGTGTGGGTGAGCAACCAACGCTGGCGCAACCAATGAGAGCGTCTTTAACTCTGA 2336
 QY 141 GlyLeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 DB 2335 GGTCTGGAAGTTCGTGAATACGCTTATGATGCGGAAATACACACTTTGACTTCGAT 2276
 QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 DB 2275 GCACGTATTAACAGCGCTGAATGAAGCTCAGCGTGGCGACGATGCTGTTCATGGCTGC 2216
 QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 DB 2215 TGCCATAACCCCAACCGGTATCGACCTACGCTGGAAACCAATGGCAACACTGGCAACATC 2156
 QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 DB 2155 TCCGTTGAGAAAGCTGTTACCGCTGTTTGACTTCGCTTACAGGGTTCGCCGCTGGT 2096
 QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 DB 2095 CTGGAAGAAGATGCTGAAGGACTGGCGCTTTTCGGCGGTATGCATAAAGAGCTGATGTT 2036
 QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluAtrqValGlyValaCysThrLeu 260
 DB 2035 GCAGTCTCTACTCTAAAAACATTTGGCCCTGTACAAGAGCGGTGGCGCTTGTACTCTG 1976
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 DB 1975 GTTGTGCGGACAGTGAACCTGTTGATCCGGATTCAGCCAATGAAAGCGCGATTCGC 1916
 QY 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValAlaThrIleLeuSerAsn 300
 DB 1915 GCTAACTACTCTAAACCCACAGCACACGCGCTTCTGTGTTGGCCACCATCTCTGAGCAAC 1856
 QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 DB 1855 GATGCGTTACGTCGATTTGGGAAACAAGAGCTGACTGATATCGCCAGCGGTATTCAGCGT 1796

QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGlySerGlyAlaAsnArgAspPheSerPhe 340
 DB 1795 ATGCGTCAGTGTTCGTCATACCTGCAGAAAAGCGCAAAACCGGACTTCAGCTTT 1736
 QY 341 IleIleLysGlnAAsnGlyMetPheSerGlyLeuThrLysGlnValLeuArg 360
 DB 1735 ATCATCAACAGAACGCGCATGTTCTCTTCAGTGGCTGACAAAGAACAAGTGTGCGT 1676
 QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 DB 1675 CTGCGCAAGAGTTTGGCGTATATCGCGTTCCTTCGTCGCTAAATGTGCGCGGATG 1616
 QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 DB 1615 ACACAGATAACATGGCTCGCTGCGAAGCGATTGTGGCAGTCTG 1568

RESULT 10
 D90730/c
 LOCUS
 DEFINITION
 D90730 AB001340
 ACCESSION
 VERSION
 D90730.1 GI:1651442
 KEYWORDS
 Complete and shotgun sequencing; msbA; ycaH; kdsB; ycbC; smtA; mukF; kicB; mukE; kicA; mukB; ycbB; aspC; ompF; tolF; cmlB; coa; cry; ams; tss.

SOURCE
 Escherichia coli K12
 ORGANISM
 Escherichia coli K12
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.

REFERENCE
 1 Oshima, T., Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Ikemoto, K., Inada, T., Itoh, T., Kajihara, M., Kanai, K., Kashimoto, K., Kimura, S., Kitagawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishio, Y., Oshima, T., Saito, N., Saito, N., Sampei, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C., Yamamoto, Y., Yano, M. and Horiuchi, T.
 A 718-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 12.7-28.0 min region on the linkage map
 DNA Res. 3 (3), 137-155 (1996)

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 2 Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiuchi, T., Ikemoto, K., Inada, T., Isono, K., Isono, S., Itoh, T., Kanai, K., Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M., Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishio, Y., Oshima, T., Saito, N., Saito, N., Sampei, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C., Yamamoto, Y. and Yano, M.
 The systematic sequencing of the Escherichia coli genome in Japan
 Unpublished
 3 (bases 1 to 20604)
 Mori, H.

TITLE
 JOURNAL
 REFERENCE
 4 Direct Submission
 Submitted (29-JUL-1996) Hirotada Mori, NARA Institute of Science and Technology, Res. & Edu. Center for Genetic Info.; 8916-5 Takayama, Ikoma, Nara 630-01, Japan
 (E-mail:hmori@gtc.aist-nara.ac.jp. Tel:81-7437-2-5660, Fax:81-7437-2-5669)
 Collaboration Information:
 Project:
 The Japan E.coli genome DNA sequencing project
 Group:
 The Japan E.coli genome DNA sequencing group
 Members: (1995.4 - 1996.3)
 Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiuchi, T., Ikemoto, K., Inada, T., Isono, K., Isono, S., Itoh, T., Kanai, K., Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M., Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishio, Y., Oshima, T., Saito, N., Saito, N., Sampei, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C., Yamamoto, Y. and Yano, M.

Headed by:

Name: Takashi Horiuchi
 Address: National Institute of Basic Biology, Okazaki, 444, Japan
 E-mail: kishori@nibb.ac.jp
 Information Operator:
 Name: Hirotada Mori
 Address: NARA Institute of Science and Technology,
 Ikoma, 630-01, Japan
 E-mail: hmori@nara.ac.jp

URL:
 The Japan E. coli genome database
 http://bsw3.aist-nara.ac.jp

FEATURES

source
 Location/Qualifiers
 1..20604
 /organism="Escherichia coli K12"
 /mol_type="genomic DNA"
 /strain="K12"
 /db_xref="taxon:83333"
 /map="20.9-21.3 min"
 /clone="Kohara clone #218"
 /notes="Nucleotide position 968194-988797 from the
 initiation site of ThrA (0 min.).
 This clone is from Kohara lambda miniset library."
 1..723
 /genes="msbA"
 <1..723
 /genes="msbA"
 /notes="ORF ID:0217#11
 similar to PIR Accession Number S27998"
 /codon_start=1
 /transl_table=11
 /product="MsbA protein."
 /protein_id="BAA35660.1"
 /db_xref="GI:1651443"
 /translation="EFRTVFTYPCGDPALRNINLKIPAGKTVALVGSQKSTIA
 SLIRFYDIEGILMDGDLREYTLASRNQVALVQNVHLFNDTVANNIAYARTEQ
 YREQTEEAARWAYAMDFINKMDGLDTVIGSGVLLSGQQRQRIARALRLDSPIL
 ILDEATSAIDTESRAIQALDELQKNTSLVIAHRLSTIEKADEIVVVEDGVIVERG
 THNDLLEHGVVYQLHKMGFGQ"
 720..1706
 /genes="ycaH"
 720..1706
 /genes="ycaH"
 /notes="ORF ID:0218#1
 similar to PIR Accession Number S27999"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein E (msbA 3' region)"
 /protein_id="BAA35661.1"
 /db_xref="GI:4062487"
 /translation="MTEKTVSGESPLWRLLLPLSWLYGLVSGAIRLCYKLLKRAWRA
 PVPVTVGNLTAGGKTPVWVWLVEQLQGRGIRVGVSRGYSYGLKAESEYPLLSADTT
 TQAGDEPVLITQRTDAPVAVSPVSDAVKAILAQHPDQIIVTDGLQRIARLADVE
 IIVIDGVRFGNGLWLPAGPMPERAGRLKSDAVIVNGVPSRSGEIPMHLPGQAVNL
 RTGRTCDVLAQLEHVAMAGIHPPRFATLKMCGVQPEKCVPLADHQSINLHADVSALV
 SAGOTLVMTKQAKVACRAFEENWYLPVDAQLSGDEPAKLLTQLTLASGN"
 1743..2975
 /notes="unnamed protein product; ORF_ID:0218#2"
 /codon_start=1
 /transl_table=11
 /protein_id="BAA35662.1"
 /db_xref="GI:4062488"
 /translation="MSLPHLSADARNHLAAQGLLNKRRRRLASLIEDIPATISRWELL
 QDITINVARSPYLVFSRLGNYPAGWDELARGELMEYWAHEACFMRSPFRLLIRH
 RLAPKEMGWKQADWAMQHEABIAQLIHIHDKGVPSRADFEHPKRGSGWGWKPKH
 KRLHGLEFTAGVWYIERNFORVYDLTRVMPDWDERDLVQSQTAEIIMLDNSNM
 LGLFROWLADYVRLKRPALAAWREARAEQQQIIIVHVEKLGNLWHDHLLPLLERAL
 AKGLTASVLSPPDPVVDKRAEQLFDFSVRLCYTPAPKROYGYFVPLLHRGQ
 LVGRMDAKHRTQGLILEVLSLWLOEGIKPTTLQKGLRQAITDFANWQATRTVTLGCC
 POGFLDCTRTGWEIDPVA"
 3027..3209
 /notes="unnamed protein product; ORF_ID:0218#3"
 /codon_start=1
 /transl_table=11

gene
 CDS
 /protein_id="BAA35663.1"
 /db_xref="GI:4062489"
 /translation="MDHRLLEIIACPVCGKGLWYNQOEKELICKLDNLNLAFLPDRDGPV
 LLEEARVLTADESKS"
 3206..3952
 /gene="kdsB"
 3206..3952
 /gene="kdsB"
 /note="ORF_ID:0218#4
 similar to PIR Accession Number A26322"
 /codon_start=1
 /transl_table=11
 /product="3-deoxy-manno-octulosonate cytidylyltransferase
 (EC 2.7.7.38)"
 /protein_id="BAA35664.1"
 /db_xref="GI:1651444"
 /translation="MSFVVIIPARYASTRLPGKPLVDINGKPMIVHVLERARESGAER
 IIVATDHEVDARVAEAGEVCMTRADHQSGERLAEVVEKCAFSDDTVIVNQGDEP
 MIPATIIQVADNLQORQVGMATLAVPHINAEARNPNAKVVLDAEGVALYPSRATI
 MPDRDFAEGLETVDGNFLRHIGIYIRAGFIRRYVNVWQPSPLEHIEMLEQLRVLMYG
 EKHVAVAEVPGTGVDTPEDLERYAEMR"
 4106..4999
 /note="unnamed protein product; ORF_ID:0218#5"
 /codon_start=1
 /transl_table=11
 /protein_id="BAA35665.1"
 /db_xref="GI:4062490"
 /translation="MEQLRAELSHLLGEKLSRIECVNEKADTALWALYDSQGNPMLP
 ARSFPGKARQLAKTMTLARSCTVRPTIYGVMTHEHPGDPDVLLERMRCVSVEA
 PARTFERQKQDQIVEALLAHRQDSRCGVADNTQENFNPWSYRQHVLEVLTILN
 QPNNTGLTMDQKRIILFRTRIECPALFEGFNDNCVLHGNFCLRSMLKDSRSDQLLAW
 GEGMLWAPREYELFRLMDNSLAEDLLMSYLQAPVPAESFIWRRMLYVLWDEVAQLVN
 TGRFRRNFPDLASKSILLPWL"
 complement(4976..5755)
 /gene="ycbC"
 complement(4976..5755)
 /gene="ycbC"
 /note="ORF_ID:0218#6
 similar to SwissProt Accession Number P36565"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein in kdsB-mukF intergenic
 region."
 /protein_id="BAA35666.1"
 /db_xref="GI:4062491"
 /translation="MLFTLKVIYGNMLPLPLMLLIIGALLWFSRFQKTGKIFIS
 IGWALLLSLQPADRLRLRPIESTYPTWNNQKVDYIVVLGGTYTNWQWAPSSNLI
 NNSLRLNEGIRLWRENGSKLIFTGVAKNTIVTAEVARGVAQSLGVPRQIITLD
 LPKDTSEAAAVKQAIQDAPFLVTSASHLPRAMIFFQOEGNLPLPAPANQLAIDSPL
 NPWRATPSVWMLHSDRVGYETLGRIMQWLKSGSGEPQRE"
 5891..6676
 /gene="smtA"
 5891..6676
 /gene="smtA"
 /note="ORF_ID:0218#7
 similar to SwissProt Accession Number P36566"
 /codon_start=1
 /transl_table=11
 /product="SmtA protein."
 /protein_id="BAA35667.1"
 /db_xref="GI:1651445"
 /translation="MQDRNFDDIAEKFSRNIYGTTKQLRQAILWQDLRVLAEWGPO
 KLRLVDAGGEGQTALKWAERGHVILCDLSAQMIDRAKQAAEAKGSVSNMOPFHCAA
 QDVASHLETFTVDLILFHAVLEWADPRSVLQTLNSVLRPGVLSLIFYNAGHLLMNM
 VAGNFDYQVAGMPKPKKKTLSPDYPRDPAQVTLWLEEGAGWIMGKTGVRVFDHLYLREK
 HQQRDCYEAALLETRYCRQEPIITLGRYIHTVARTKPSQSKDY"
 6673..7995
 /gene="kicB"
 /note="synonym: mukF"
 6673..7995
 /gene="kicB"
 /note="ORF_ID:0218#8"

Alignment Scores:

Pred. No.: 2.09e-147 Length: 20604
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservativeness: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-10-673-786A-2 (1-396) x D90730 (1-20604)

Qy	1	MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe	20
Db	18063	ATGTTTGAGAACATTACCCGCCCTCTGCCGACCCGATTCTGGCGCTGGCCCATCTGTTT	18004
Qy	21	ArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValTyrLysAspGluThr	40
Db	18003	CGTGGCGATGACGTCGCCGCAAAATTAACCTCGGGATTGGTGTCTATAAGATGAGAG	17944
Qy	41	GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu	60
Db	17943	GGCAAAACCCCGGTACTGACACAGCGTGAAGAGGCTGAACAGTATCTGCTCGAAAATGAA	17884
Qy	61	ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu	80
Db	17883	ACCACCAAAATTAACCTCGGCATTGACGCGCATCCCTGGAATTTGGTGGCTGCACACTCAGGAA	17824
Qy	81	LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr	100
Db	17823	CTGCTGTTTGGTAAGGTAGCGCCCTGATCAATGACAAACGTGCTCGACGGCACAGACT	17764
Qy	101	ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal	120
Db	17763	CGCGGGGGCACGTGGCGCATACGCTGGCTGGCGATTTCTGGCAAAAATACACAGCGTT	17704
Qy	121	LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla	140
Db	17703	AAGCGGTGTGGGTGAGCAACCAACAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA	17644
Qy	141	GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAenHisThrLeuAspPheAsp	160
Db	17643	GGTCTGGAAATTCGTGAATACGCTTATATGATGGGAAATACACACTCTTGACTTCGAT	17584
Qy	161	AlaLeuIleAsnSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys	180
Db	17583	GCACCTGATTACAGCGCTGAATGAAGCTCAGGCTGGCGAGTGTGCTGTTCCATGGCTGC	17524
Qy	181	CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu	200
Db	17523	TGCCATAACCCCAACCGGTATGACCTACGCTGGAAACCAATGGCAACACTGGCACAACTC	17464
Qy	201	SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly	220
Db	17463	TCCGTGTGAGAAAGGTGTGTACCGTGTGTGACTTCGTTACCGAGGTTTGGCCGCTGGT	17404
Qy	221	LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaIleMetHisLysLeuLeuVal	240
Db	17403	CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCCGGCTATGCAATAAGAGCTGATTGTT	17344
Qy	241	AlaSerSerTyrSerLysAsnPheGlyLeuTyrAenGluArgValGlyAlaCysThrLeu	260
Db	17343	GCCAGTTCCTACTCTAAAACATTTGGCCCTGTACAAACGAGCGGTGTGGCGCTTGTACTCTG	17284
Qy	261	ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg	280
Db	17283	GTGTCTGGCAGTGAACACCGTTGATCCCGATTACGCCAATGAAGCGCGATTTCGC	17224
Qy	281	AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn	300
Db	17223	GCTAACTACTCTAACCCACACAGCACCGCGCTTCTGTTGTTGCCACCATCTGTAGCAAC	17164
Qy	301	AspAlaLeuArgAlaIleTyrGluGlnLeuLeuThrAspMetArgGlnArgIleGlnArg	320
Db	17163	GATCGGTTCACGTGCGATTGGGAACAAGAGCTGACTGATATATGCGCCAGCGTATTACGCGT	17104

Qy	321	MetArgGlnLeuPheValAsnThrLeuGlnGlnLysGlyAlaAsnArgAspPheSerPhe	340
Db	17103	ATGGCTCAGTTGTTTCGTCAATACGCTGAGGAAAAAGCGCAAAACCGGACTTCAGCTTT	17044
Qy	341	IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg	360
Db	17043	ATCATCAAAACAGAACGGCATGTTCTCTTCAGTGGCCCTGACAAAAGAACAGTGTCTGCT	16984
Qy	361	LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet	380
Db	16983	CTGGCGAGAGATTGGCGTATATGCGGTGCTTCTGCTCGGTAATGTGGCCCGGATG	16924
Qy	381	ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu	396
Db	16923	ACACAGATAACATGGCTCCGCTGTGCGAAGCGATTGTGGCAGTGTCTG	16876

RESULT 11

U00096_09/c

WPCOMMENT

Sequence split into 47 fragments LOCUS U00096 Accession U00096

Fragment Name	Begin	End
U00096_00	1	110000
U00096_01	100001	210000
U00096_02	200001	310000
U00096_03	300001	410000
U00096_04	400001	510000
U00096_05	500001	610000
U00096_06	600001	710000
U00096_07	700001	810000
U00096_08	800001	910000
U00096_09	900001	1010000
U00096_10	1000001	1110000
U00096_11	1100001	1210000
U00096_12	1200001	1310000
U00096_13	1300001	1410000
U00096_14	1400001	1510000
U00096_15	1500001	1610000
U00096_16	1600001	1710000
U00096_17	1700001	1810000
U00096_18	1800001	1910000
U00096_19	1900001	2010000
U00096_20	2000001	2110000
U00096_21	2100001	2210000
U00096_22	2200001	2310000
U00096_23	2300001	2410000
U00096_24	2400001	2510000
U00096_25	2500001	2610000
U00096_26	2600001	2710000
U00096_27	2700001	2810000
U00096_28	2800001	2910000
U00096_29	2900001	3010000
U00096_30	3000001	3110000
U00096_31	3100001	3210000
U00096_32	3200001	3310000
U00096_33	3300001	3410000
U00096_34	3400001	3510000
U00096_35	3500001	3610000
U00096_36	3600001	3710000
U00096_37	3700001	3810000
U00096_38	3800001	3910000
U00096_39	3900001	4010000
U00096_40	4000001	4110000
U00096_41	4100001	4210000
U00096_42	4200001	4310000
U00096_43	4300001	4410000
U00096_44	4400001	4510000
U00096_45	4500001	4610000
U00096_46	4600001	4639675

Continuation (10 of 47) of U00096 from base 900001 (U00096 Escherichia coli K-12 MG1655

Alignment Scores:

Pred. No.: 1.53e-146 Length: 110000

Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-10-673-786A-2 (1-396) x U00096_09 (1-110000)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 84932 ATGTTTGAGAACATTACCGCGCTCTGCGGACCCGATTCTGGGCTGGCGATCTGTTT 84873
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyLeuValIleGlyValIleAspGluThr 40
Db 84872 CGTGGCGATGAACGTCGCGGCAAAATTAACTCGGGATTGGTGTCTATAAAGATGAGACG 84813
Qy 41 GlyIleThrProValLeuThrSerValIleLysAlaGluGlnIleLeuGluAsnGlu 60
Db 84812 GGCAAAACCCCGGTACTACCGAGCGTGAAAGAGCTGAACAGTATCTGCTCGAAATGAA 84753
Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 84752 ACCACCAAAATTACTCGCGATTGACGGCATCCCTGGAATTTGGTCGCTGCACTCAGAA 84693
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 84692 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGTCGCGACGACAGACT 84633
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 84632 CCGGGGGCACTGGCGCACTACGCGTGGTGGCGATTTCTGGCAAAAATACACGCGTT 84573
Qy 121 LysArgValTropValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 84572 AAGCGTGTGGTGAGCAACCAAGCTGGCGCAACCAAGAGCGTCTTTAACTCTGCA 84513
Qy 141 GlyLeuGluValArgGlyTyrAlaTyrTrpAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 84512 GGTCTGGAAGTTCGTGAATACGCTATTATGATGCGGAATATCACACTCTTGACTTCGAT 84453
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 84452 GCACGTATTAAACAGCTCAATGAAGCTCAGGCTGGCGAGCTAGTCTGTTCCATGGCTGC 84393
Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTropGlnThrLeuAlaGlnLeu 200
Db 84392 TGCCATAACCAACCGGTATCGACCTACGCTGGAAACCAATGGCAACACTGGCAACTC 84333
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 84332 TCCGTTGAGAAGGCTGTTACCGCTGTTTGCCTTACCGAGGTTTGGCGGTGT 84273
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 84272 CTGGAAGAAGATGCTGAAGGACTGGCGCTTTCGGCGGTATGCATAAAGAGCTGATGTT 84213
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 84212 GCCAGTTCTTACTCTAAAACTTTGGCGCTGTACACAGAGCGTGTGGCGCTGTGACTCTG 84153
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 84152 GTTGTGCGGACAGTGAACCGTTGATCGCGATTACGCCAAATGAAAGCGGCGATTGCG 84093
Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 84092 GCTAACTACTCTAACCCACGACACACGCGCTTCTGTTGTTGCCACCATCTGAGCAAC 84033
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnIleLeuThrAspMetArgGlnArgIleGlnArg 320
Db 84032 GATCGTTTACGTGGATTTGGAAACAAGAGCTGACTGATATGCGCGACGCTATTGAGCGT 83973
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnIleLysGlyAlaAsnArgAspPheSerPhe 340

Db 83972 ATGCGTCAGTTGTTGTCGTAATACGTCGAGAAAAGCGCAAAACCGGACTTCAGCTTT 83913
Qy 341 IleIleLysGlnAsnGlyMetPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 83912 ATCATCAAAAGAACGCGCATGTTCTCTTCAGTGGCCCTGACAAAAGAACAGTGTGCGT 83853
Qy 361 LeuArgGluGluPheGlyValIleAlaValAlaLaserGlyArgValAsnValAlaGlyMet 380
Db 83852 CTGCGCGAAGAGTTTGGCGTATATGCGGTTCCTTCGTGCGTAAATGTCGCGGATG 83793
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 83792 ACACCAATAACATGCTCGCTGTCGGAAGCGATTGTGGCAGTCTGTG 83745
RESULT 12
AE005174_11/c
WPCOMMENT
Sequence split into 56 fragments LOCUS AE005174 Accession AE005174
Fragment Name Begin End
AE005174_00 1 110000
AE005174_01 100001 210000
AE005174_02 200001 310000
AE005174_03 300001 410000
AE005174_04 400001 510000
AE005174_05 500001 610000
AE005174_06 600001 710000
AE005174_07 700001 810000
AE005174_08 800001 910000
AE005174_09 900001 1010000
AE005174_10 1000001 1110000
AE005174_11 1100001 1210000
AE005174_12 1200001 1310000
AE005174_13 1300001 1410000
AE005174_14 1400001 1510000
AE005174_15 1500001 1610000
AE005174_16 1600001 1710000
AE005174_17 1700001 1810000
AE005174_18 1800001 1910000
AE005174_19 1900001 2010000
AE005174_20 2000001 2110000
AE005174_21 2100001 2210000
AE005174_22 2200001 2310000
AE005174_23 2300001 2410000
AE005174_24 2400001 2510000
AE005174_25 2500001 2610000
AE005174_26 2600001 2710000
AE005174_27 2700001 2810000
AE005174_28 2800001 2910000
AE005174_29 2900001 3010000
AE005174_30 3000001 3110000
AE005174_31 3100001 3210000
AE005174_32 3200001 3310000
AE005174_33 3300001 3410000
AE005174_34 3400001 3510000
AE005174_35 3500001 3610000
AE005174_36 3600001 3710000
AE005174_37 3700001 3810000
AE005174_38 3800001 3910000
AE005174_39 3900001 4010000
AE005174_40 4000001 4110000
AE005174_41 4100001 4210000
AE005174_42 4200001 4310000
AE005174_43 4300001 4410000
AE005174_44 4400001 4510000
AE005174_45 4500001 4610000
AE005174_46 4600001 4710000
AE005174_47 4700001 4810000
AE005174_48 4800001 4910000
AE005174_49 4900001 5010000
AE005174_50 5000001 5110000
AE005174_51 5100001 5210000
AE005174_52 5200001 5310000

AE005174_53	5300001	5410000
AE005174_54	5400001	5510000
AE005174_55	5500001	5528445

Continuation 712 of 56) of AE005174 from base 1100001 (AE005174 Escherichia coli O157:H7

Alignment Scores:

Pred. No.:	6,348-146	Length:	1100000
Score:	2037.00	Matches:	3394
Percent Similarity:	99.7%	Conservative:	1
Best Local Similarity:	99.5%	Mismatches:	1
Query Match:	99.6%	Indels:	0
DB:	1	Gaps:	0

US-10-673-786A-2 (1-396) x AE005174 11 (1-110000)

Qy	1	MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe	20
Db	104403	ATGTTTGAGAACATTTACCGCGCTCCTCCGACCGAATTCGGGCTGGCCGATCTGTTT	104344
Qy	21	ArgAlaAspGluArgProGlyIysIleAsnLeuGlyIleGlyValTyrIysAspGluThr	40
Db	104343	CGTGCCGATGAACGTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG	104284
Qy	41	GlyLysThrProValLeuThrSerValIysLysAlaGluGlnTyrLeuLeuGluAsnGlu	60
Db	104283	GGCAAAACCCCGGTACTGACCAGCGTGAAAAGGCTGAACAGTATCTGCTGCAAAATGAA	104224
Qy	61	ThrThrIysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu	80
Db	104223	ACCACCAAAATTAACCTCGGATTTGACGGCATCCCTGAATTTGGTTCGTGCTACTCAGGAA	104164
Qy	81	LeuLeuPheGlyLysGlySerAlaIleuIleAsnAspLysArgAlaArgThrAlaGlnThr	100
Db	104163	CTGCTGTTTGTGAAGGTAGCGCCCTGTATCAATGCAACCGTCTCGCAGCGCACAGACT	104104
Qy	101	ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal	120
Db	104103	CCGGTGGCACTGGCGCACTACGCATAGCTGCCGATTTCTTGGCAAAAATAACACAGCTT	104044
Qy	121	LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla	140
Db	104043	AAGCGAGTGTGGGTGAGCAACCCCAAGCTGGCGCAACCATAGAAGCGTCTTTAACTCTGCA	103984
Qy	141	GlyLeuGluValArgGluTyrAlaTyrThrAspAlaGluAsnHisThrLeuAspPheAsp	160
Db	103983	GATCTGGAAGTTCGTGAAATACGCTTATTATGATCGGAAAAACCAACCCCTTGACTTCGAT	103924
Qy	161	AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys	180
Db	103923	GCAC TGATTAAACAGCTGAACGAAGCTCAGCGTGGCAGCTAGTGTCTTCCATGGCTGC	103864
Qy	181	CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu	200
Db	103863	TGCCACAACCCCAACCGGTATCGACCCCTACGCTGGAAACAATGGCAGACACTGGCAACAATC	103804
Qy	201	SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly	220
Db	103803	TCCGTTGAGAAAGCTGGTATACCGCTGTTTTCGACTTCGCTTACCGGGTTTTGCCCGTGGT	103744
Qy	221	LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal	240
Db	103743	CTGGAAGAGNATGCTGAAGGACTGGCGCTTTCGGGCTATGTCATAAGAGCTGATTTGTT	103684
Qy	241	AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu	260
Db	103683	GCCAGTTCCCTACTCTAAAAATTTTGGCCTGTACACAGCGCTGTTGGCGCTTGTACTCTG	103624
Qy	261	ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg	280
Db	103623	GTTGCTGCCGACACTGAAACCGTTGATCGCGATTCACCCAAATGAAGACGCGCATTCGC	103564
Qy	281	AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn	300

```
AE005174_46 4600001 4710000
AE005174_47 4700001 4810000
AE005174_48 4800001 4910000
AE005174_49 4900001 5010000
AE005174_50 5000001 5110000
AE005174_51 5100001 5210000
AE005174_52 5200001 5310000
AE005174_53 5300001 5410000
AE005174_54 5400001 5510000
AE005174_55 5500001 5528445
Continuation (13 of 56) of AE005174 from base 1200001 (AE005174 Escherichia coli O157:H7)

Alignment Scores:
Pred. No.: 6,34e-146 Length: 110000
Score: 2037.00 Matches: 394
Percent Similarity: 99.7% Conservativity: 1
Best Local Similarity: 99.5% Mismatches: 0
Query Match: 99.6% Indels: 0
DB: 1 Gaps: 0

US-10-673-786a-2 (1-396) x AE005174_12 (1-110000)

QY 1 MetPheGluAsnIleThrAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
DB 4403 ATGTTTGAGAACATTACCGCGCTCTGCGCGACCGAATTCGGCGCTGCGCATCTGTTT 4344
QY 21 ArgAlaAspGluArgProGlyLysIleAanLeuGlyIleGlyValTyrLysAspGluThr 40
DB 4343 CGTGGCGATGAACGTCCTCCGGCAAAATTAACCTCGGGATTTGGTGCTATTAAGATGAGCG 4284
QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
DB 4283 GGCACAAACCCGGTAGTACCGAGCGTGAAAGAGCTGAACAGTATCTGCTCGAAATGAA 4224
QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
DB 4223 ACCACCAAAATTAACCTCGGCATTCACGCGCATCCCTGGAATTTGGTGGCTGCACACTCAGGAA 4164
QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAanAspLysArgAlaArgThrAlaGlnThr 100
DB 4163 CTGCTGTTTGGTAAGGTAGCGCCCTGATCAATGACAAACGTCGCGACGGCAGACT 4104
QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
DB 4103 CCGGGTGGCACTGGCGCACTAGCATAGCTGCGATTCCTGCGCAAAATACACGCTT 4044
QY 121 LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
DB 4043 AAGCGAGTGGGTGAGCAACCAAGCTGGCGCAACCAATAGAGCGCTCTTTAACTCTGCA 3984
QY 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
DB 3983 GATCTGGAAAGTTCGTAATACCTTATTATGATGCGGAACCAACACACCCCTTGACTTCGAT 3924
QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
DB 3923 GCACCTGATTACAGCTGAACAGCTCAGCTAGCTGGCGACGTAAGTCTGTTCCATGGCTGC 3864
QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
DB 3863 TGCCACAACCCACCGGTATCGACCTAGCTGGGAACAAATGCGACACTGCGCACAACTC 3804
QY 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
DB 3803 TCCGTTGAGAAAGCTGGTTACCGCTGTTTGACTTCGCTTACGAGGTTTTCGCCGTGTGT 3744
QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
DB 3743 CTGGAAGAGATGCTGAAGGACTGCGCGCTTTTCGGCGCTATGCATAAAGAGCTGATGTT 3684
QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValCylValaCysThrLeu 260
DB 3683 GCCAGTTCTCTATAAAACCTTTGGCGCTGTACACGAGCGTGTGGCGCTTGTACTCTG 3624
```

RESULT 14

BA000007_11/c

WPCOMMENT

Sequence split into 55 fragments LOCUS BA000007 Accession BA000007

Fragment Name	Begin	End
BA000007_00	1	110000
BA000007_01	100001	210000
BA000007_02	200001	310000
BA000007_03	300001	410000
BA000007_04	400001	510000
BA000007_05	500001	610000
BA000007_06	600001	710000
BA000007_07	700001	810000
BA000007_08	800001	910000
BA000007_09	900001	1010000
BA000007_10	1000001	1110000
BA000007_11	1100001	1210000
BA000007_12	1200001	1310000
BA000007_13	1300001	1410000
BA000007_14	1400001	1510000
BA000007_15	1500001	1610000
BA000007_16	1600001	1710000
BA000007_17	1700001	1810000
BA000007_18	1800001	1910000
BA000007_19	1900001	2010000
BA000007_20	2000001	2110000
BA000007_21	2100001	2210000
BA000007_22	2200001	2310000
BA000007_23	2300001	2410000
BA000007_24	2400001	2510000
BA000007_25	2500001	2610000
BA000007_26	2600001	2710000
BA000007_27	2700001	2810000
BA000007_28	2800001	2910000
BA000007_29	2900001	3010000
BA000007_30	3000001	3110000
BA000007_31	3100001	3210000
BA000007_32	3200001	3310000
BA000007_33	3300001	3410000
BA000007_34	3400001	3510000
BA000007_35	3500001	3610000
BA000007_36	3600001	3710000
BA000007_37	3700001	3810000
BA000007_38	3800001	3910000


```

/gene="dacC"
/locus_tag="c0924"
/EC_number="3.4.16.4"
/function="enzyme; Murein sacculus, peptidoglycan"
/notes="D-alanyl-D-alanine carboxypeptidase fraction C;
Escherichia coli K-12 ortholog: b0839; Escherichia coli
O157:H7 ortholog: z1066"
/codon_start=1
/transl_table=11
/product="Penicillin-binding protein 6 precursor"
/protein_id="AAN79397.1"
/db_xref="GI:26107214"
/translacion="MDTRVAFMTQYSSILRLGAAGSAFLFAPTAFAAQTVEAPSV
DAWATLMDYAGSKVLAEGNADEKLDPAKTLKIMTSYVVGAKADKIKLDMVTVGK
DAWATGSSVNFVLPKPGQVADLNKGVIITQSGNDACIALADVAGSQESFIFG
LMNGYAKGLGNTTFTVTHGLDAPQGFSTARDMALLGKALIHDPPEYVATHEKEFT
FNKIQPNRLLWSSNLNVDMKGTGTAGAGYNLVASATQDMRLISVVLGAKTDRI
RNSEKLLTWGFRFEFTPIPKPATFTVQKRWFGDKSEVNLGAGEAGSVTIPIRGQL
KNLKASYLTLPQLPKKGQVGTIDFQLNGKSIQRPLIVMENVEEGFFGRMWID
FVMKFFHQWFGSWFS"
/complement(1411..2169)
/gene="deor"
/locus_tag="c0925"
/complement(1411..2169)
/gene="deor"
/locus_tag="c0925"
/function="regulator; Salvage of nucleosides and
nucleotides"
/notes="Escherichia coli K-12 ortholog: b0840; Escherichia
coli O157:H7 ortholog: z1067"
/codon_start=1
/transl_table=11
/product="Deoxyribose operon repressor"
/protein_id="AAN79398.1"
/db_xref="GI:26107215"
/translacion="METREERIEGQLLQELKRSKLHLKDAALGVSEMTIRDLNN
HSAPVLLGGYTVLPEPRASHYLLDQKSRVVEERAKAALATVBDQTLFFDCGT
TPTVIEADNEIPTAVCYSLNTFLAKEKPHCFACGGEFHASNAIFKPIDFOQT
LNNFCPDIAFYSAAGVHYSKGCATCNLELPVKHWMSMAQKHLVVDHSFEGKVRPA
RMGLDKRFDIVSDCCPEDEYKYAQTQRIKLMY"
/complement(2227..2823)
/gene="ybjG"
/locus_tag="c0926"
/complement(2227..2823)
/gene="ybjG"
/locus_tag="c0926"
/notes="Escherichia coli K-12 ortholog: b0841; Escherichia
coli O157:H7 ortholog: z1068"
/codon_start=1
/transl_table=11
/product="Hypothetical protein ybjG"
/protein_id="AAN79399.1"
/db_xref="GI:26107216"
/translacion="MLENLNLSFLSINATPDSAPWMISLAIFIAKDLITVVPILAAV
LWLWGLVITFEALAFICWRLHSGSLMLVLAWSRVYLVGHVWDLMLGGLLAGMIG
CLSAQIIWQAGCHKLYQRLQSWYRCPALPIRKWVRD"
/3078..4340
/gene="cmr"
/locus_tag="c0927"
/3078..4340
/gene="cmr"
/locus_tag="c0927"
/function="transport; Transport of small molecules: Other"
/notes="Chloramphenicol resistance pump cmr; Escherichia
coli K-12 ortholog: b0842; Escherichia coli O157:H7
ortholog: z1069"
/codon_start=1
/transl_table=11
/product="Multidrug translocase mdfA"
/protein_id="AAN79400.1"
/db_xref="GI:26107217"
/translacion="NRDYWRNCNQKSLASGRQALLPPLCLVLYEFTYIGND

```

```

MIQCMILAVVEQYQAGIDWVTSMTAYLAGCMFLQWLLGCLPLSDRIGRRPVMLAGVWF
IVTCLAILAQNIQFLLRLOQISLCFICAVGAYAAIQESFEPAVCIRIKITALMANVA
LIAPLLGPLVGAAMIHVLPWEGMFVFAALAAISFPGLQRAMPETATIGEKLSKREL
GRDYKVLKNGRFVAGALAGFVSLPIAWIAQSPFIIITGEQLSSYSGVGLQVPIFG
ALAGNLLARLTSSRTVRSLLIIMGWPIIMIGLVAATAATVISSHAYLWMTAGLSIYA
FGIAGNAGLVRLTLFASDMSKGTVSAMGMLQMLIFTVVGIEISKHAMLINGGGLPNL
FNLVNGILWLSLWIFLKDQMGNSHEG"
/complement(4381..4665)
/gene="ybjH"
/locus_tag="c0928"
/complement(4381..4665)
/gene="ybjH"
/locus_tag="c0928"
/notes="Escherichia coli K-12 ortholog: b0843; Escherichia
coli O157:H7 ortholog: z1070"
/transl_table=11
/product="Hypothetical protein ybjH precursor"
/protein_id="AAN79401.1"
/db_xref="GI:26107218"
/translacion="WIMKNCILLGALLMGFTGVMAQSVTVVPSGYKVVVVPDSVSV
POAVSVATPOTVVAAPAPAPAYRPHVVRHLASVGEGMVIEHQIDDDHH"
/complement(4751..5569)
/locus_tag="c0929"
/complement(4751..5569)
/locus_tag="c0929"
/notes="Escherichia coli K-12 ortholog: b0844; Escherichia
coli O157:H7 ortholog: z1071"
/codon_start=1
/transl_table=11
/product="Protein ybjI"
/protein_id="AAN79402.1"
/db_xref="GI:26107219"
/translacion="WMSIKLIAVDMDGTFLSDQTYNREEREMAOYQOMKAQGIKRVVA
SGNQYQIISFPPELANEIAFVAENGHWVSEKDVFNGLSKDAPATVVEHLLTRPE
VEIACGKNSAYTLKKYDDAMTKVAEMTYHRLVYVDNFNDLEIDFFKGLNLSDELIP
QVQALHEAIGDIMVPHVTGMSIDLIIIPGVHKNGLRQLQKLGIDSEVVVFDGGG
NDIEMLRQAGSFAMENAGSAVAAAKYRAGSNREGVLDVIDKVLKHEAPFKQ"
/complement(5566..6837)
/gene="ybjJ"
/locus_tag="c0930"
/complement(5566..6837)
/gene="ybjJ"
/locus_tag="c0930"
/function="putative regulator"
/notes="Escherichia coli K-12 ortholog: b0845; Escherichia
coli O157:H7 ortholog: z1072"
/codon_start=1
/transl_table=11
/product="Hypothetical protein ybjJ"
/protein_id="AAN79403.1"
/db_xref="GI:26107220"
/translacion="MCTNVSLHNAVVPPLLITWMTSVNSRNALKRRTWALFMPFFFL
PGLMLASWATTPAIRDILSVISAEMGVLEGLSIGSMGILCSAWLVKRGTRNVIL
LPMHGFYSLGTLAGAGVMALTAFGVPATVHLLAALVGIAPVIAIQAIPIDGTGKN
AADTQGEKGVPEYRDIQLLIGVWVLAFAESANDMLPMLWDGHEGFSPTSSSL
IYAGTGLTGVTRFIFGWFIDHYSRVAVFRASALMGALGILIIIFVDSANWAGVSVL
WGLGASLGFPLTISAADTPDAPTRVSVVATGYLAFVLGPPLLGLGEGHYGLRSAM
LVVLALVILAAIVAKAVAKPDTKTQTAMENS"
/6828..7394
/gene="ybjK"
/locus_tag="c0931"
/6828..7394
/gene="ybjK"
/locus_tag="c0931"
/function="putative regulator"
/notes="Escherichia coli K-12 ortholog: b0846; Escherichia
coli O157:H7 ortholog: z1073"
/codon_start=1
/transl_table=11
/product="Hypothetical protein ybjK"
/protein_id="AAN79404.1"

```

```
/db_xref="GI:26107221"
/translation="MYTIAERGNMRRANDPORREKLIQATLEAVKLYGIHAVTHRKI
AASGVATPDNMELMYQYALASRKPLIKTVQNMQRSSQOTLEQMFEGTARALDAFI
EGMTLHFVTDKPLSRBEILRMVERIAG"
gene
complement(7873..8904)
/feature="intn"
/locus_tag="c0932"
complement(7873..8904)
/feature="intn"
/locus_tag="c0932"
/function="enzyme; Integration, recombination Phase or
Prophage Related"
/notes="Escherichia coli O157:H7 ortholog: z2966"
/codon_start=1
/transl_table=11
/product="Integrase for prophage"
/protein_id="AAN79405.1"

Alignment Scores:
Pred. No.: 4,27e-145 Length: 301276
Score: 2033.00 Matches: 392
Percent Similarity: 100.0% Conservative: 4
Best Local Similarity: 99.0% Mismatches: 0
Query Match: 99.4% Indels: 0
DB: 1 Gaps: 0

US-10-673-786A-2 (1-396) x AE016758 (1-301276)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 135038 ATGTTTGAGAACATTACCGCGCTCTGCCACCGGATTCTGGGCGCTGGCGATCTGTTT 134979

Qy 21 ArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValTyrLysAspGluThr 40
Db 134978 CGTCCGATGAACGTCGCGCAAAATTAACCTCGGATTTGTTGTTCTATAAGATGAGACG 134919

Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 134918 GGTAAACCCCGGTACTGACCCAGCGTGAAAGGCTGAAACAGTATCTGCTCGAAATGAA 134859

Qy 61 ThrThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 134858 ACCACCAAAATTAACCTCGCATTTGACCGCTCCCTGAAATTTGGTCGCTGACCTCAGAA 134799

Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
Db 134798 CTGCTGTTTGGTAAGTAGGCTGCTGATCAATGACAAACGCTGCTCGACCGGACAGACT 134739

Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 134738 CCGGTGGCACTGGCGCACTACGCGTGCTGCCGATTTCTTGGCAAAATAACACGCTT 134679

Qy 121 LysArgValTrpValSerAsnProSerTrpProAenHisLysSerValPheAsnSerAla 140
Db 134678 AAGCGTGTGTGGGTGAGCAACCAAGCTGGCGCAACCAATATAGAGCGCTTTAACTTCGA 134619

Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAspPheAsp 160
Db 134618 GGTCTGGAAGTTCGTGAATACGCTTATATGATCGGAAATACACACTCTTGACTTCGAT 134559

Qy 161 AlaLeuIleAenSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 134558 GCACGTATTACAGCTGAACGAGCTCAGGCTGGCGAGTAGTGTGTTCCATGGGTGC 134499

Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 134498 TGCCATAACCAACCGCATCGACCTACGCTGGAAACAAATGGCAGACACTGGCACAACCTC 134439

Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 134438 TCCGTTTGAAAGGCTGGTTTACCGCTGTTTACCTTCGCTTACCAAGGGTTTGGCCCGTGT 134379
```

```
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 134378 CTGAAGAAGATGCTGAGGACCTGCGGCTTTCCGGCTCTGCATAGAGGAGCTGATGTT 134319

Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 134318 GCCAGTTCTTACTCTAAAACCTTTGGCTCTGTACAAACGAGCGTGTGGCGCTTGTACTCTG 134259

Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 134258 GTTGCTCTGACAGTGAACCTGTTTGAACGCGCATTCAGCCAAATGAAAGCGCGCATTCGT 134199

Qy 281 AlaAenTyrSerAenProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 134198 GCTAACTACTCTAACCCACCAGCACACGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 134139

Qy 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 134138 GATGCGTTTACGTGGATTTCGGAAACAAGAACTGACTGATATGCGCCAGCGTATTACGCGT 134079

Qy 321 MetArgGlnLeuPheValAenThrLeuGlnGluLysGlyAlaAenArgAspPheSerPhe 340
Db 134078 ATGCGTCAGTTGTTGCTCAATACGCTGACGAGAAAGGCGCAAAATCGCGACTTCAGCTTT 134019

Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 134018 ATCATCAACAGAACCGCATGTTCTCTCTCAGTGGCTGTGACAAAGAACCAAGTCTGCGT 133959

Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenValAlaGlyMet 380
Db 133958 CTGGCCCAAGAGTTTGGCGTATATGCTGTTGCTTCTGCTCGCTAAACGTAGCTGGGATG 133899

Qy 381 ThrProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 133898 ACACCAATTAACATGGCGCGCTGTGCAAGCGATTGTGCGCAGTGTCTG 133851
```

Search completed: March 15, 2006, 16:49:38
Job time : 5484 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2006, 02:39:56 ; Search time 447 Seconds
(without alignments)
4736.188 Million cell updates/sec

Title: US-10-673-786A-1

Perfect score: 1191

Sequence: 1 atgttgagaacattaccgc.....cgattggcagtgctgttaa 1191

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:**

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/PCFUS COMB.seq.*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1191	100.0	1331	3	US-08-985-908-23
2	839	70.4	1209	3	US-09-489-039A-4902
3	458.2	38.5	1830121	3	US-09-557-884-1
4	458.2	38.5	1830121	3	US-09-643-990A-1
5	458.2	38.5	1830121	3	US-10-158-865-1
C 6	387.6	32.5	579	3	US-09-489-039A-4954
7	348.4	29.3	702	3	US-09-543-681A-3805
C 8	316.4	26.6	1509	3	US-09-543-681A-12968
9	316.4	26.7	1632	3	US-09-543-681A-12507
C 10	306.6	25.7	24417	2	US-08-846-762-1
11	305	25.6	1242	3	US-09-252-991A-12428
C 12	305	25.6	1329	3	US-09-252-991A-12162
13	289	24.3	1338	3	US-09-489-039A-1301
C 14	278.2	23.4	1185	3	US-09-489-039A-1274
15	276.4	23.2	1209	3	US-09-540-236-315
16	274.8	23.1	100848	3	US-09-596-002-39
17	263.6	22.1	1284	3	US-09-328-352-769
18	242.6	20.4	1260	3	US-09-543-681A-3123
19	218.8	18.4	2440	3	US-09-976-594-429
20	201.2	16.9	426	3	US-09-543-681A-3717
21	177.8	14.9	1941	3	US-09-220-132-65
22	177.8	14.9	1941	3	US-09-814-915A-35
23	177.8	14.9	1942	3	US-09-949-016-1216
24	176.2	14.8	2056	3	US-09-566-921-56

Sequence 59, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 4873, Ap
Sequence 1412, Ap
Sequence 12380, A
Sequence 145, App
Sequence 13293, A
Sequence 3374, Ap
Sequence 4336, Ap
Sequence 3375, Ap
Sequence 12333, A
Sequence 1235, Ap
Sequence 5, Appli
Sequence 658, App
Sequence 4632, Ap
Sequence 316, App
Sequence 952, App
Sequence 3376, Ap
Sequence 1603, Ap
Sequence 434, App

25 154.4 13.0 1127 3 US-09-533-559-59
26 151.2 12.7 1230025 3 US-09-198-452A-1
27 151.2 12.7 1230230 3 US-09-438-185A-1
28 143.6 12.1 207 3 US-09-489-039A-4873
29 142.2 11.9 900 3 US-09-533-559-4412
30 125.4 10.5 423 3 US-09-252-991A-12380
31 117.4 9.9 988 3 US-09-533-559-145
32 116.4 9.8 564 3 US-09-252-991A-13293
33 107.2 9.0 912 3 US-09-248-796A-3374
34 103.2 8.7 668 3 US-09-533-559-4436
35 96.6 8.1 867 3 US-09-248-796A-3375
36 90.2 7.6 423 3 US-09-252-991A-12333
37 89.6 7.5 306 3 US-09-489-039A-12335
38 79 6.6 1808 3 US-09-646-925-5
39 70.6 5.9 646 3 US-09-328-111-658
40 69.4 5.8 658 3 US-09-533-559-4632
41 63.8 5.4 279 3 US-09-540-236-318
42 61.4 5.2 500 3 US-09-533-559-952
43 60 5.0 567 3 US-09-248-796A-3376
44 53.2 4.5 264 3 US-09-313-294A-1603
45 53.2 4.5 417 3 US-09-328-111-434

ALIGNMENTS

RESULT 1
US-08-985-908-23
; Sequence 23, Application US/08985908
; Patent No. 6004773
; GENERAL INFORMATION:
; APPLICANT: MASAYUKI ARAKI, MASAKAZU SUGIMOTO, YASUHIKO YOSHIHARA, AND TSUYOSHI NAKA
; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/985,908
; FILING DATE: 05-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-325659
; FILING DATE: 05-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: NORMAN F. OBLON
; REGISTRATION NUMBER: 24,618
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1331 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; STRAIN: JM109
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10..1197
US-08-985-908-23

```
Query Match      100.0%; Score 1191; DB 3; Length 1331;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTGAGAAATACCGCGCTCTGCGCGACCCCGATTCTGGCCCTGGCCGATCTGTTT 60
DB 10 ATGTTTGAGAAATACCGCGCTCTGCGCGACCCCGATTCTGGCCCTGGCCGATCTGTTT 69
QY 61 CTGCGCGATGAACGTCCCGCAAAATTAACCTCGGATTTGTTCTATAAAGATGAGACG 120
DB 70 CTGCGCGATGAACGTCCCGCAAAATTAACCTCGGATTTGTTCTATAAAGATGAGACG 129
QY 121 GCAAAACCCCGTACTGACACGGCTGAAAAGGCTGAACAGTATCTGTCGAAAATGAA 180
DB 130 GCAAAACCCCGTACTGACACGGCTGAAAAGGCTGAACAGTATCTGTCGAAAATGAA 189
QY 181 ACCACCAAAATTAACCTCGGCATTTGACCGCATCCCTGAATTTGGTCTGCTCACTCAGGAA 240
DB 190 ACCACCAAAATTAACCTCGGCATTTGACCGCATCCCTGAATTTGGTCTGCTCACTCAGGAA 249
QY 241 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGTCGCGCAACACAGACT 300
DB 250 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGTCGCGCAACACAGACT 309
QY 301 CCGGGGGGCACTGGCGCACTACGGTGGCTGCCGATTTCTGCGCAAAATACACGGTT 360
DB 310 CCGGGGGGCACTGGCGCACTACGGTGGCTGCCGATTTCTGCGCAAAATACACGGTT 369
QY 361 AAGCGTGTGGTGAGCAACCAACGCTGGCCGCAACCAATGAAGCGCTTTAACTCTGCA 420
DB 370 AAGCGTGTGGTGAGCAACCAACGCTGGCCGCAACCAATGAAGCGCTTTAACTCTGCA 429
QY 421 GGTCTGGAAGTTCGTGAATACGCTTATATGATCGGAAATCACACTTTGACTTCGAT 480
DB 430 GGTCTGGAAGTTCGTGAATACGCTTATATGATCGGAAATCACACTTTGACTTCGAT 489
QY 481 GCACTGATTAACAGCCTGAATGAAGCTGAGCTGGCGACGTAGTCTGTTCCATGGCTGC 540
DB 490 GCACTGATTAACAGCCTGAATGAAGCTGAGCTGGCGACGTAGTCTGTTCCATGGCTGC 549
QY 541 TGGCATAAACCAACCGGTATCGACCTTACGCTGGAAACAACTGGCAACACTGGCACAACTC 600
DB 550 TGGCATAAACCAACCGGTATCGACCTTACGCTGGAAACAACTGGCAACACTGGCACAACTC 609
QY 601 TCCGTTGAGAAAGCTGGTTACGCTGTTTGAATTCGCTTACAGGGTTTGGCCGCTGCT 660
DB 610 TCCGTTGAGAAAGCTGGTTACGCTGTTTGAATTCGCTTACAGGGTTTGGCCGCTGCT 669
QY 661 CTGGAAGAAGATGCTGAAGACTGCGCGCTTTTCGCGCTATGCAATGAAGACTGATTGTT 720
DB 670 CTGGAAGAAGATGCTGAAGACTGCGCGCTTTTCGCGCTATGCAATGAAGACTGATTGTT 729
QY 721 GCCAGTTCTCTACTTAATAAATTTGGCTCTGTAACAGAGCGTGTGGCGCTTGTACTCTG 780
DB 730 GCCAGTTCTCTACTTAATAAATTTGGCTCTGTAACAGAGCGTGTGGCGCTTGTACTCTG 789
QY 781 GTTGCTGCCGACAGTGAACCGTTGATPCGCGCATTCAGCCAAATGAAGCGGCGATTGCG 840
DB 790 GTTGCTGCCGACAGTGAACCGTTGATPCGCGCATTCAGCCAAATGAAGCGGCGATTGCG 849
QY 841 GCTAACTACTCTAACCCACAGACACAGGCGCTTCTGTTGTCACCATCTCTGAGCAAC 900
DB 850 GCTAACTACTCTAACCCACAGACACAGGCGCTTCTGTTGTCACCATCTCTGAGCAAC 909
QY 901 GATGCGTTAGTCGATTTGGAACACAGAGCTGACTGATATGCGCCAGCGTATTTACGCGT 960
DB 910 GATGCGTTAGTCGATTTGGAACACAGAGCTGACTGATATGCGCCAGCGTATTTACGCGT 969
QY 961 ATGCGTCAGTTGTTGCTCAATAGCTGCAGGAAAAAGCGCAAAACCGCGACTTTGCTTTT 1020
DB 970 ATGCGTCAGTTGTTGCTCAATAGCTGCAGGAAAAAGCGCAAAACCGCGACTTTGCTTTT 1029
QY 1021 ATCATCAACAGAACGGCATGTTCTCTTTCAGTGGCTCTGACAAAGAACAAAGTGTGCGT 1080
```

```
DB 1030 ATCATCAACAGAACGGCATGTTCTCTTCACTGCGCTGACAAAAGAACAAAGTGTGCGT 1089
QY 1081 CTGCGCGAAGAGTTTGGCGTATATGCGGTGCTTCTGTCGCGTAATGTCGCCGGATG 1140
DB 1090 CTGCGCGAAGAGTTTGGCGTATATGCGGTGCTTCTGTCGCGTAATGTCGCCGGATG 1149
QY 1141 ACACAGATAAATGCTCGCTGCGGAGGATTTGTGCGAGTCTGTAA 1191
DB 1150 ACACAGATAAATGCTCGCTGCGGAGGATTTGTGCGAGTCTGTAA 1200

RESULT 2
US-09-489-039A-4902
; Sequence 4902, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4902
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4902

Query Match      70.4%; Score 839; DB 3; Length 1209;
Best Local Similarity 81.5%; Pred. No. 6.7e-273;
Matches 971; Conservative 0; Mismatches 220; Indels 0; Gaps 0;

QY 1 ATGTTTGAGAAATACCGCGCTCTGCGCGACCCCGATTCTGGCCCTGGCCGATCTGTTT 60
DB 19 ATGTTTGAGAAATACCGCGCTCTGCGCGACCCCGATTCTGGCCCTGGCCGATCTGTTT 78
QY 61 CTGCGCGATGAACGTCCCGCAAAATTAACCTCGGATTTGGTCTTATAAAGATGAGACG 120
DB 79 CTGCGCGATGAACGTCCCGCAAAATTAACCTCGGATTTGGTCTTATAAAGATGAGACG 138
QY 121 GCAAAACCCCGTACTGACAGCGTGAAGAGGCTGAACAGTATCTGCTGCAAAATGAA 180
DB 139 GGTAAACCGCGCTTCTGACAGCGTCAAAAAAGCAGACGATCTGCTGGAATAATGAA 198
QY 181 ACCACCAAAATTAACCTCGGCATTTGACGCGATCCCTGAATTTGGTCTGCTCACTCAGGAA 240
DB 199 AGACTTAAAACTATCTGGGCATCGATGGTATTCCTGAAATTTGGTCTGCTCACTCAGGAG 258
QY 241 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGTCCTGCGACGGCAACAGACT 300
DB 259 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGCGTGAATAAACGCGCCGCGCAGACA 318
QY 301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCCGATTTCTTGGCAAAATACACAGGTT 360
DB 319 CCGGGGGGCACTGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 378
QY 361 AAGCGTGTGGTGAGCAACCCAGCTGCGCAACCAATGAAGCGTCTTTAACTCTGCA 420
DB 379 AAACGTGTGGGTGAAGTAAATCCGAGCTGGCGCAACCAATGAAGCGTATTTACCTCTGCC 438
QY 421 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATACCACTTTGACTTCGAT 480
DB 439 GGGCTGGAAGTGGCGCAATACGCATACGACGCGGCTAACCCACGCGCTGGACTTTGAT 498
QY 481 GCACTGAATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTCTGTTCCATGGCTGC 540
DB 499 GGTCTGCTGCCAGCGCTGAACGAAAGCCAGCGCGGCGACGTGGTACTGTTTCCACGCGTGC 558
QY 541 TGCATAAACCCACCGGTATCGACCTTCGATCGCTGGAACAATGGCAAAACACTGGCACACTC 600
```

Db 559 TGCCACACCCGACGCGTATCGATCCGACGCTCGATCAGTGGCAGCAGCTGCGCAGCTG 618
Qy 601 TCCGTTGAGAAAGGCTGTTACCGCTGTTTACCTTTCGCTTACCGAGGTTTTCGCGGCTG 660
Db 619 TCGGTGAAAAGGCTGCTACCGCTGTTTCGATTTTCGCTTACCGAGGCTTTCGCGCGGT 678
Qy 661 CTGGAAGAAGTGTGAAGSACTGCGCGCTTTTCGCGGCTATGCAATAAGAGCTGATTGTT 720
Db 679 CTGGAAGAAGTGTGAAGGCTGACGCGCTTTTGTCTCTACATAAAGAGCTGCTGTC 738
Qy 721 GCCAGTTCTACTCTAAAACTTTGGCTGTACACGAGCGTGTGGCGCTTGTACTCTG 780
Db 739 GCCAGTCTCTACTGAAAACTTTGGCTGTACACGAGCGCTGCGCGCTGCACTCTG 798
Qy 781 GTTGTGCGCAGTGAACCGTTGATCGCGCATTTCAAGCCAAATGAAGCGCGCATTCGC 840
Db 799 GTGCGCGCGATCAGGAGCTGTAGACCGCGCTTCAGTCAGATGAAGTCGGTATCGC 858
Qy 841 GCTAACTACTTAACCCACAGCACACGCGCTTCTGTTGTGGCCACCATCTCAGCAAC 900
Db 859 GCCAACTACTGAAACCGCTGCGATGCGCGCTCCGTTGTCGCCACCATTTCTGAGCAAC 918
Qy 901 GATCGTTACTGCGATTGTTGGAAACAAGAGCTGATGATATGCGCGCAGCGTATTCAGCGT 960
Db 919 GATCGCTTACGCGCAATCTGGGAGCAGGAATGACCGATATGCGCCAGCGCATCCAGCGT 978
Qy 961 ATGCGTCAAGTTGCTCAATACCTGCGAGGAAAAGCGCAACCGCGACTTCAGCTTTT 1020
Db 979 ATGCGTCTGCTGTTGCTCAATACCTGCGAGGAAAAGCGCGAGCGGACTTCAGCTTT 1038
Qy 1021 ATCATCAACAGACGCGCATGTTCTCTTCAGTGGCTTGACAAAGAACAAAGTCTGCGT 1080
Db 1039 ATCAGCAGCAGACGCGCATGTTCTCATTCAGCGGCTTGACTAAAGCAGCTGCTGCGC 1098
Qy 1081 CTGCGGAAAGTTTGGCGTATATGCGTTGCTTCTGTCGCGTAAATGTGCGCGGATG 1140
Db 1099 CTGCGTGAAGATTGCGCATCTATGCGTAGCTTTCGAGACGTATCAACGTTGCGCGGATG 1158
Qy 1141 ACACGAGATACATGCTCGCTGCGAAGCGATTGCGCAGTCTGTAA 1191
Db 1159 ACGCTTGACAAATATGGCGCGCTGTGGAAGCCATCGTCGCGCTACTGTAA 1209

RESULT 3
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match 38.5%; Score 458.2; DB 3; Length 1830121;
Best Local Similarity 61.5%; Pred. No. 5.5e-142;
Matches 733; Conservative 0; Mismatches 458; Indels 0; Gaps 0;

Qy 1 ATGTTTTCAGAAACATTACCGCGCTCTCGCGACCCGATTTCTGGGCGCTTGGCGGATCTGTTT 60
Db 1684147 ATGTTTTCAGAAACATTCAAGCGCGCACCGATCTCAATCTTAGGCTTAGGCGAAGCAATT 1684206
Qy 61 CGTCCGATGAACGTCCCGGCAAAATTAACCTCGGATTCGGTGTCTATAAAGATGAGACG 120
Db 1684207 AAATCCGAAACTCGCGAAATAAAATCAATTTGGGTATTGGCGTTTATAAAGATGCGCAA 1684266
Qy 121 GGCMAAACCCGGTACTGACCGGTTGAAAAGGCTGAAAAGCTGAAACAGTATCTGCTCGAAATGAA 180
Db 1684267 GGCACACACCCCAATTATGACGCGGTAAAAGAGCCGAAACGATATTATTGATAAGGAA 1684326
Qy 181 ACCACCAAAATTAACCTTCGCGCATTTGACGGCATCCCTGAAATTTGGTCTGCTGACCTCAGGAA 240
Db 1684327 AAAACCAAGAAATATCTGACTATCGATGTTTGGGATTTATACGAAACAAACAAAGCA 1684386
Qy 241 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGTGCTCGCAGCGCACAGACT 300
Db 1684387 CTCCTTTTCGTTAAAGATTCTGAAGTCTCAATCTAATCGAGCAAGAACAGTACAAAGT 1684446
Qy 301 CCGGGGGGCACTGCGGCACTACGGTGGCTGCGGATTTCTTCGCAAAATAATACCAGCGTT 360
Db 1684447 TTAGCGCGAAACAGGTGCATTTACGCATTTGCGCGAGAAATTTATTAACGCCAACTTAAAGCA 1684506
Qy 361 AAGCGTGTGTTGAGCAACCCAGCTGCGCAACCATAGACGCTCTTTAACTCTGCA 420
Db 1684507 CAAATGTTTGGWTTCAGCACACCACTTTGCCAAACACACATGCGATTTTCAATGCTGTC 1684566
Qy 421 GGTCTGGAAGTTCGTGAATACGCTTATATGATCGCGGAAATCACACTTTGACTTTCGAT 480
Db 1684567 GGTATGACCATTCGTGAATATCGTTATATGATGCTGAAACGCAAGGCCCTTTGATTGGGAA 1684626
Qy 481 GCATGATTAACAGCTGAATGAAGTCAAGCTGCGGAGCGAGTGTGCTTCCATGCGTGC 540
Db 1684627 CACTTATTAGAAGATTTAAGCCAAAGCAAGCGGATGTGTTGCTTTTACACGGTTGT 1684686
Qy 541 TGCCTAATCCCAACCGGTATCGACCTTACCTCGAACAATGGCAAAACACTTGGCAACAATC 600
Db 1684687 TGCCTAATCCGACTGTTGTTGCTTACCTTCCAGAAACAAATAGCCGCACTT 1684746
Qy 601 TCCGTTGAGAAAGGCTGTTTACCGCTGTTTTCGCTTTCACAGGCGTATTGCGCGGTGTT 660
Db 1684747 TCAGCTAATAATGTTGTTGCTTTCGCACTTTTTCGCTTATCAAGTTTAGCCAAACGGA 1684806
Qy 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGGCTATGCAATAAAGAGCTGATTGTT 720
Db 1684807 TTAGATGAAGATGCTTATGTTTACGTCGCTTTTTCGAGCAAAACCAACAAAGAAATATTAGT 1684866
Qy 721 GCCAGTTCTACTCTAATAAATTTTGGCTGTACACGAGCGTGTGCGCTTGTACTCTG 780
Db 1684867 GCGAGTTCTATTCTCGAAAAAATTTTGGTTTATATAAATGAACGTTGTGTCATTTACCCCTT 1684926
Qy 781 GTTGTGCGCAGTGAACCGGTTGATCGCGCATTCAGCCAAATGAAGAGCGGCGATTTCGC 840

```
Db 1684927 GTGCGAAAAATGCAGAAATTGCAATCAACTATTAAACCCAGTAAATAATCAATTATTCG 1684986
Qy 841 GCTAACTACTCTAACCCACAGACACAGCGGCTTCTGTTGTCGCCACCATCTCGAGCAAC 900
Db 1684987 ACCCTACTCTAACCCAGCTTCTCACAGGGGGCGACCGTAGCAACAGTATTAAATGAT 1685046
Qy 901 GATCGGTTAGCTGCGATTGCGGAACAAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT 960
Db 1685047 GCTCAACTTGGCCAGAAATGGGAAATGAATTAATTAATGAAATGGTGAAACGATCAAAAA 1685106
Qy 961 ATGCGTCAGTTGTTTCGTCATACGCTGCGAGAAAAAGCGCAAAACCGCGACTTTCAGCTTT 1020
Db 1685107 ATGCGTCACATTATTCGTTTCAGTTATTAAAGAAATATGTCGAGAAACAAGATTTCAGCTTT 1685166
Qy 1021 ATCATCAACAGAACGCGCATGTTCTCTTCAGTGGCCGTGACAAAGAACAAAGTCTGCGT 1080
Db 1685167 ATCATTAACAAACCGTATGTTTCTTCAGTGGATTAAACAGGGGAAACAAGTGGATCGT 1685226
Qy 1081 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGCGTAAATGTCGCGCGGATG 1140
Db 1685227 TTAATAAATGAAATTTGCGCATTTAGCTGTTCTGTCGCTATCAACGTAAGTGAATC 1685286
Qy 1141 ACACAGATAACATGGCTCCGCTGTCGGAAGCGATTGTGGCAGTGTCTGTAA 1191
Db 1685287 ACAGAATAATATTCCTATCATATGTAAGTATCGTGAAGTACTTTAA 1685337
```

RESULT 4

```
US-09-643-990A-1
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter
;
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
;
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB186P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
```

```
;
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match      38.5%; Score 458.2; DB 3; Length 1830121;
Best Local Similarity 61.5%; Pred. No. 5.5e-142;
Matches 733; Conservative 0; Mismatches 458; Indels 0; Gaps 0;

Qy 1 ATGTTTGAGAAACATTACGCGCTCTGCGGACCCGAGTATCTGGGCTCTGGCCGATCTGTTT 60
Db 1684147 ATGTTTGAAACATATCAAGCGCGCACCGAGTCCAAATCTTAGGCTTAGCGAAGCATTT 1684206
Qy 61 CGTGCGGATGAACGTCGCCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 120
Db 1684207 AAATCGGAACCTCGCGAAATAAATAATCAATTTGGGTATTGGCGTTTATAAAGATGCGCA 1684266
Qy 121 GGCMAAACCCCGGTACTGACACGCGTGAATAAAGGCTGAACAGATATCTGCTCGAAAAATGAA 180
Db 1684267 GGCACAAACCCCAATTATGACACGCGGTAAAGAAGCCGAAACAGGATTAATTGATAAGGA 1684326
Qy 181 ACCACCAAAATTAATCTCGGCATTTGACGGCATCCCTGAAATTTGGTCGCTGCACTCAGGAA 240
Db 1684327 AAAACCAAGAATTATCTGACTATCGATGTTATTCGGGATTATAACGAAACAAACAAAGCA 1684386
Qy 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCCTCGCACGGCACAGACT 300
Db 1684387 CTCCTTTTCGGTAAAGATTCTGAAAGTCATCCAATCTAATCGAGCAAGAACAGTACAAAGT 1684446
Qy 301 CGGCGGGGCACTGGCGCACTAGCGTGGCTGCGGATTTCTTGGCAAAAAAATACACGCGTT 360
Db 1684447 TTAGCGGAAACAGGTGCATTACGCATTGGCGGAGATTTATTAAAGCCCAACATAAGCA 1684506
Qy 361 AAGCGTGTGTGGGTGAGCAACCCAAAGCTGGCCGAAACCAATAAGAGCGTCTTTAACTCTGCA 420
Db 1684507 CAAAATGTTTGGWTGAGCACACCAACTTTGGCCAAACCAACAATGCGGATTTTCAATGCTGC 1684566
Qy 421 GGTCTGGAAGTTTCGTGAATACGCTTATTATGATGCGGAAATACACACTCTTGACTTCGAT 480
Db 1684567 GGTATGACCAATTCGTGAATATCGTTATTATGATGCTGAAACGAAAGCCCTTGTATTGGGAA 1684626
Qy 481 GCACATGATTAAACAGCTGAATGAAGCTCAGGCTGGCGAGTGTAGTGTCTTCAATGCTGTC 540
Db 1684627 CACTTATTAGAAGATTAAAGCAAGCAAGCGAGGCGATGTGTGCTTTTACACGTTGT 1684686
Qy 541 TGCCATAAACCCAAACCGGTATCGACCTCTAGCTGGAACAAATGGAACAACTGGGCACAACTC 600
Db 1684687 TGCCATAATCCGACTGGTATTGACCCCTACTCCAGAAACAATGGCAAGAAATTAGCCGCACTT 1684746
Qy 601 TCGGTTGAGAAAGGCTGGTTACCGCTGTTTGACTTCGCTTACGAGGTTTTCGCCGTGGT 660
Db 1684747 TCAGCTTAAATAATGGTTGGTTGCCACTCTTTTGACTTTTGTCTTATCAAGGTTTAGCCACGGA 1684806
Qy 661 CTGGAAGAAGATGCTGAAGGAGCTGGCGCTTTTCGGCGCTATGTCATAAAGAGCTGATTGTT 720
Db 1684807 TTAGATGAAGATGCTTATGTTTACGTCCTTTTTCGAGCAAAACCAAGAAATATTAGTG 1684866
Qy 721 GCCAGTTCTCTACTCTAAAAAATTTTGGCCCTGTGTCAACGAGCGTGTGGCGCTTTGACTCTG 780
Db 1684867 GCGAGTTTCATTCTCGNAAAACCTTTGGTTTATATAATGAACGTTGTGTCGATTTACCCCTT 1684926
Qy 781 GTTGCTGCCGACAGTGAACCGGTTGATTCGGGATTCAGGCAATGAAGCGGGGATTCGC 840
Db 1684927 GTGGCAGAAAATGCAGAAATTCGATCAACCTCATTAACCAAGTAAATAATCAATTATTCGC 1684986
Qy 841 GCTAACTACTTAACCCACAGCACACGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 900
Db 1684987 ACCCTATACTCTAACCCAGCTTCTCACGGCGGGGCGACCGTAGCAACAGTATTAAATGAT 1685046
Qy 901 GATGCGTTACGTGCGATTTTGGGAAACAAAGAGCTGACTGATATGCGCCAGCGGTATTCAGCGT 960
Db 1685047 GCTCAACTTCGCCAAGAAATGGGAAATGAATTAACCTGAAATGCGTGAACGCAATCAAAAA 1685106
```


; NAME/KEY: misc_feature		Query Match	38.5%; Score 458.2; DB 3; Length 1830121;
		Best Local Similarity	61.5%; Pred. No. 5.5e-142;
		Matches	733; Conservative 0; Mismatches 458; Indels 0; Gaps 0;
QY	1	ATGTTTGAGAACATTACCGCGCTCTCGCGACCCCGATTCTGGCGCTTGGCGCTGCGCGATCTGTTT	60
DB	1684147	ATGTTTGAACATATCAAAGCGGCACGCGATCCCAATCTTAGCTTTAGCGAAGCATTT	1684206
QY	61	CGTCCGATGAACGTCCCGCAAAATTAACCTCGGATTGGTCTCTATAAAGATGAGACG	120
DB	1684207	AAATCCGAAACTCGGAAATAAATCAATTTGGGTATTGGCGTTTATAAAGATGCGCAA	1684266
QY	121	GGCAAAACCCCGTACTGACAGCGTGAAAGGCTGAACAGTATCTGCTCGAAATGAA	180
DB	1684267	GGCAACCCCAATTAATGCACGCGTAAAGAGCGCAAAACGATTATTGTATAAGGAA	1684326
QY	181	ACCAACCAAAATTTACTCTGGCATTTAGCGGATCCCTGAATTTGGTGCCTGCACTCAGGA	240
DB	1684327	AAAACCAAGAAATTAATCTGACTATCGATGGTATTGGCGATTATACGAAACAAACGA	1684386
QY	241	CTGCTGTTTGTAAAGGTAGCGCCTGATCAATGACAAACGTCTCGCACGGCACAGACT	300
DB	1684387	CTCCTTTTCGGTAAAGATTCTGAAGTCAATCCAATCTAATCGACAGAACAGTACAAAGT	1684446
QY	301	CGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTCGCAAAATAATCACGCGTT	360
DB	1684447	TTAGCGGACAGGTGCAATACGCAATTCGCGCAGAAATTTATTAAACGCCAACTAAAGCA	1684506
QY	361	AAGCGTGTGGGTGAGCAACCCCAAGCTGCGCGAAACCAATGAGAGCTCTTTAACTCTGCA	420
DB	1684507	CAAAATGTTTGGWTGAGCACACCAACTTGGCCAAACCCACAATCGGATTTTCAATGCTGC	1684566
QY	421	GCTCTGGAAGTTCGTGAATACGTTTATATGATGCGGAAATCACACTCTTGACTTCGAT	480
DB	1684567	GGTATGACCAATTCGTGAATATCGTTATTATGATGCTGACGAAAGCCCTTGTATGGAA	1684626
QY	481	GCACCTGATTAAACAGCTGAATGAAGCTCAGGCTGGCGACGATGCTGCTGTTTCCATGCTGC	540
DB	1684627	CACTTATTAGAGATTTAAGCCAAGCAAGCAAGGCGATGCTGCTTTTACACGGTTGT	1684686
QY	541	TGCCATAACCAACCGGTATCGACCTTACGCTGGAACAATGGCAACACACTGGCACACTC	600
DB	1684687	TGCCATAATCCGACTGGTATTGACCTTACTCCAGAACAAATGGCAAGAAATTAGCCGACTT	1684746
QY	601	TCCGTTGAGAAAGCTGTTTACCGCTGTTTGACTTCGCTTACCAGGCTTTTGCCCGTGT	660
DB	1684747	TCAGCTAAAAATGGTGGTGCCACTCTTTGCACTCTTTGCACTTTCTATCAAGTTTAGCCACGG	1684806
QY	661	CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTGCGGCTATGATCAAGAGCTGATTGTT	720
DB	1684807	TTAGATGAGATGCTTATGTTTACGTGCTTTTGCAGCAACCAACCAAGAAATATTAGTG	1684866
QY	721	GCCAGTTCCTACTCTAAAAACTTTGGCCTGTACAAAGAGGCTGTTGGCGCTTGTACTCTG	780
DB	1684867	GCGAGTTCATTCTCGAAAACTTTGGTATTATAATGAACAGTGTGTGTGATTTTACCCTT	1684926
QY	781	GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCAAAATGAAAGCGCGATTCGC	840
DB	1684927	GTGGCAGAAAAATGCAGAAATTTGCATCAACCTTAAACCCCAAGTAAATCAATTTATTCG	1684986
QY	841	GCTAACTACTCTAACCCACAGCACGCGGCTTCTGTTGTGCAACCATCTCGACCAAC	900
DB	1684987	ACCCTATACTCTAACCCAGCTTCTCACGGGGGGCGACCGTAGCAACAGTATTAAATGAT	1685046
QY	901	GATGCGTTACGTGCGATTTTGGGAAACAAGAGCTGACTGATATGCGCCACGCTATTACGCT	960
DB	1685047	GCTCAACTTCGCCAAGAATGGAAAAATGAATTAATCAATGCTGAACGATCAAAAA	1685106
QY	961	ATGCGTCAAGTTGTTGCTCAATACGCTCGAGAAAAAGCGCAACCGGACTTTCAGCTTT	1020
DB	1685107	ATGCGTCACTTATTCGTTTCAGTTATTAAGAATATATGCAAGCAAGATTTTTCAGCTTT	1685166

QY 661 CT 662
 Db 697 CT 698

RESULT 8
 US-09-252-991A-12968/c
 ; Sequence 12968, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 12968
 ; LENGTH: 1509
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-12968

Query Match 26.6%; Score 316.4; DB 3; Length 1509;
 Best Local Similarity 54.9%; Pred. No. 5.6e-96;
 Matches 645; Conservative 0; Mismatches 526; Indels 3; Gaps 1;

QY 4 TTTGAGAACATTACCGCGCTCTCGCGACCCGATCTGGGCTTGGCCGATCTGTTTGGT 63
 Db 1290 TTGCGAAGTGTGCGCGGTACCGGCGACCCGATCTGGGCTTGGCCGATCTGTTTGGT 63

QY 64 GCGGATGAAGTCCCGCAAAATTAACTCGGAGTGGTGTCTATTAAGATGAGCGGC 123
 Db 1230 AACGATCCGCGCGGCAAGCTGGACCTCGGCGTGGTGTCTACAAAGGATGCCAGGGC 1171

QY 124 AAAACCCGCTACTGACGACGCTGAAAGGCTGAAACAGTATCTGCTCGAATAATGAAC 183
 Db 1170 CTGACCCGATCTGCGCTCGGTGAACTCGCGGACGCGCTGCTGAGCAGGAAC 1111

QY 184 ACCAAATTAATCTCGGATGACGCGATCCCTGAATTTGGTCTGCTGCACTCAGGAATG 243
 Db 1110 ACCAAGAGTACGTGCGCGGCGACCGGATGCGCTGTTGCGCGCGCTTGGCGGAATG 1051

QY 244 CTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAAGCTGCTCGCAGCGGACAGACTCG 303
 Db 1050 GCGCTCGGCGCGCTCGCGCTGTTGCTGGAGCAAGCGCGCGACCCAGCGCC 991

QY 304 GGGGGCACTGGCGACTACGCGTGGCTGCGGATTTCTGGCAAAATAATACAGCTTAAG 363
 Db 990 GGGGGCACTGGCGCTGCGCTGGCGGAGCTTATCGCCATTCGCTGCGCGCGCG 931

QY 364 CGTGTGTGGTGAGCAACCCAAAGCTGGCGAACCAATGAAGCGCTTTAACTCTGAGGT 423
 Db 930 GGCATCTGGCTGAGCGACCCGACCTGGCCGATCCACGAGACCTGTTTGGCGCGCGCG 871

QY 424 CTGGAAGTTCTGTAATAGCTTATATGATGCGGAAATCACACTTTGACTTCGATGCA 483
 Db 870 CTGAAGGTTTCCCACTACCCCTAGCTCAGCGCGCGACCAACCGC---CTGGATTCGAGGCG 814

QY 484 CTGATTAACAGCTTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTTCCATGCTGCTGC 543
 Db 813 ATGCTTCTGCTTGGAGCGCATTCGCCAGGAGACGTGTGTGCTGCTGATGCTGCTGC 754

QY 544 CATAACCAACCGGATTCGACCTTACGCTGGAACAATGGAACAACACTGGCACAACCTTCC 603
 Db 753 CACAACCGACCGGTTTCGACCTGAGCCACGACGACTGGCGAGGGGTGCTCGACGTGGT 694

QY 604 GTTGAGAAAGGCTGTTACCGCTGTTTGAATTCGCTTACCAGGGTTTGGCCGCTGCTG 663

Db 693 CGTGGCGGAGCTGCTGCGCGCTGATCGACTTTCGCTACGAGGCTTCGCGGACGCTCTC 634
 QY 664 GAAGAAGATGCTGAAGGACTGCGCGCTTTCGCGGCTATCATTAAGAGAGTGTGTTGCC 723
 Db 633 GAGGAAGACGCTGCGCGGTACGCTTTCGCGCGGCAACTGCCGAGGTGCTGGGTCAACC 574
 QY 724 AGTTCCTACTCTAAAAAATTTTGGCCTGTACAAAGAGCGGTGTTGGCGCTTGTACTCTGGTT 783
 Db 573 AGTTCCTGCTCGAAGAACTTCGCGCTTGTACCGGACCGCGTTCGGGCGCTGATCGTCTGC 514
 QY 784 GGTGCGGACAGTGAACCCGTTGATCGCGATTCAGCCAAATGAAGCGGCGATTCGCGCT 843
 Db 513 GCGCAGAACCGCAGAGCTCACGACCTGCGTAGCAACTGCGCTTTCCTCGGCCGAAAC 454
 QY 844 AACTACTTAACCCACCGACGACACGCGGCTTCTGTTGTTGGCCACCATCTCTGAGCAACGAT 903
 Db 453 CTCTGTGTCGACCCCGCGCGCATGGCGCGAGGTGTCGCGCGCATCTCTCGCGACAGC 394
 QY 904 GGTTCAGTGGGATTTGGGAAACAAGAGCTGACTGATATGCGCAGCGTATTCAGCGTATG 963
 Db 393 GAGTTGAAGGACTTTGGCAGGAAGAGTTCGAAGGATGCGCTCGCGCATCGCCAGCCTG 334
 QY 964 CGTCAAGTTTCTGTCATACGCTGCGAGAAAGGCGCAAAACCGCGACTTTCAGCTTTATC 1023
 Db 333 CGCATCGGCTGTCGAAGCCCTGCGCGCGCAGCGCTGCGCGAGGCTTCGCCCATGTC 274
 QY 1024 ATCAAAACAGAACGGCATTTCTCTTCAGTGGCGCTGACAAAAGAACAGTGTGCGCTG 1083
 Db 273 GCGCGCAACGCGGGATGTTTCTCTATACCGGACTGAGCCCGCAGCAGGTGCTCGGCTG 214
 QY 1084 CGCGAGAGTTTGGCGTATATGCGTGTCTCTGTCGCGTAAATGTCGCCGCGATGACA 1143
 Db 213 CGCGAGAGCAGCGCGCTTTACCTGTGTCCAGCGCGCGGCGCAACGTCGCCGCTATAGAC 154
 QY 1144 CCAGATAACATGCTCGCTGCGGAAAGCGATTG 1177
 Db 153 GCGCGCGCTGCGCGCTGCGCAAGCCATCG 120

RESULT 9
 US-09-252-991A-12507
 ; Sequence 12507, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 12507
 ; LENGTH: 1632
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-12507

Query Match 26.6%; Score 316.4; DB 3; Length 1632;
 Best Local Similarity 54.9%; Pred. No. 5.9e-96;
 Matches 645; Conservative 0; Mismatches 526; Indels 3; Gaps 1;

QY 4 TTTGACAACATTACCGCGCTCTCGCGACCCGATTCGGGCTTGGCCGCTGCTGTTTCTG 63
 Db 442 TTGCGCAAGGTGCGCGCGGTACCGGCGACCCGATCTCGGCGCTGCTCGACGCTACCGC 501

QY 64 GCGGATGAAGCTCCCGCGCAAAATTAACCTCGGAGTTGGTGTCTATTAAGATGAGACGGC 123
 Db 502 AACGATCCGCGCGGACAGCTGGACCTCGGCGTGGTGTCTTACAGGATGCCAGGGC 561

```

Qy 124 AAAACCCCGGTACTGACAGCGCTGAAAGAGGTGAACAGTATCTGCTGAAATGAAC 183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
562 CTGACCCCGACTCGTGGCTCGTGAACCTGCCGAGCAGCCCTGGTCGACAGAAAC 621
Qy 184 ACCAAATATCTCGGCAITGACGGCAATCCCTGAATTTGCTGCTGCACTCAGAACTG 243
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
622 ACCAAGACTACGTCCGCGGCAACGGCGATCGCTGTTTCGCGCGCGCTGCGGAACTG 681
Qy 244 CTGTTTGGTAAAGTAGCGGCTGATCATGACAAAGTGTCTGCAACGGCAGACTCCG 303
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
682 GCGCTCGCGCGCGCTCGCGCTGTTGCTGAGCAACGCGCGGACGCCACCGACGCC 741
Qy 304 GGGGGCACTGGCGCACTACGGTGGCTGCCGATTTCTTGGCAAAATACCAAGGTTAAG 363
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
742 GCGGACACGGCGCTTGGCGCTGGCGGCACTTCATCGGCCATTCGCTGCCGCGCG 801
Qy 364 CGTGTGTGGTGAGCAACCCAAAGTGGCGCAACCAATAGAGCGTCTTTAACTCTGAGGT 423
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
802 GGCATCTGGCTGAGCGACCCGACCTGGCCGATCCACGAGACCTGTTTCGCGCGCGCG 861
Qy 424 CTGGAAGTTCGTGAATAGCTTATATGATCGGAAATACACTCTTGACTTCGATGCA 483
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
862 CTGAAGGTTTCCCACTACCCCTACGTCAGCGCCGACCAACCGC---CTGGATGTCGAGCG 918
Qy 484 CTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGGCTGTG 543
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
919 ATGCTTGTGCTGGAGCGCATTCGCCGAGGAGACGTGTGCTGCTGCACTGCTGTG 978
Qy 544 CATAACCCAAACCGGTATCGACCTTACGCTGGAACAAATGGCAACACTGGCAACTCTCC 603
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
979 CACAACCCGACCGCTTCGACTCAGCCACGACGACTGGCGACGGGTCTCGAGGTGTG 1038
Qy 604 GTTGAGAAAGCTGTTACGCTGTTGACTTCGCTTACAGGGTTTTTGGCGGTGTGCTG 663
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1039 CGTCCCGCAGAGTGTGCGCTGATCGACTTCGCTTACCAGGGCTTCGGCGACGGTCTC 1098
Qy 664 GAAGAAGTGTGAAGGACTCGCGCTTTTGGCGCTATGCTAAGAGCTGATTTGTTGCC 723
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1099 GAGGAAGCGCTGGGGGTGACGCTGTTCCCGGGGACATGCGCGAGGTGCTGGTCAAC 1158
Qy 724 AGTTCCTACTATAAACTTTGGCTGTACAAAGAGGTGTGCGCTGTGACTCTGCTGTT 783
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1159 AGTTCCTGCTGAAGAACTTCGGCTGTACCGGACCGCGTTCGGGGGCTGATGCTGTC 1218
Qy 784 GCTGCCGACAGTGAACCGTTGTATCGGCGATTCAGCCAAATGAAGCGGGATTCGCGCT 843
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1219 GCGCAGAACCGCGAAGACTCACCGACCTGCTAGCCAACTGGGCTTCTCGCCCGAAAC 1278
Qy 844 AACTACTTAACCCACGACGACGCGCTTCTGTTGTCACCATCTCTGAGCAAGAT 903
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1279 CTCGTGTCAGACCCCGCGCGCATGCGCGAGGTGTCGCGGGATCTCTCGCGGACAGC 1338
Qy 904 GCGTTACGTGCAATTTGGGAACAGAGCTGACTGATATGGCCAGCGTATTCAGCGTATG 963
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1339 GAGTTGAAGGACTTTGGCAGGAGAGGTGCAAGGCATGCGCTCGCGCATCGCCAGCTG 1398
Qy 964 GTCAGTTGTTGTCATATGCTGCAAGAAAGGCGCAACCGCGCACTTCAGCTTTATC 1023
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1399 CGCATCGGCTGTGTCGAAGCGCTGGCGCGCACGCGCTGGCGGAGCTTCGCGCATGTC 1458
Qy 1024 ATCAACAGAAACCGCATGTTCTCTTCAAGTGGCTGCAAAAGAAAGTGTCTGCTG 1083
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1459 GCGCGCAACGCGGGAATGTTTCTTATACCGGATGAGCGCGCAGCAGGTGCTCGGCTG 1518
Qy 1084 CCGGAAGAGTTTGGCGTATATGCGGTGCTTCTGGTCCGCTAAATGTTGGCGCGGATGACA 1143
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1519 CGCGACGAGCAGCGCTTTACCTGTTCTCAGCGCGCGCGGCAACGTCGCGGTATAGAC 1578
Qy 1144 CCAGATTAACATGGCTCGCTGTGCGAAGCGATTG 1177
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1579 GCGCGCGCTCGACCGCTGGCGAAGCATCG 1612

```

```

RESULT 10
US-08-846-762-1/c
; Sequence 1, Application US/08846762A
; Patent No. 5994072
; GENERAL INFORMATION:
; APPLICANT: Lam, Joseph S.
; APPLICANT: Burrows, Lori
; APPLICANT: Charlier, Deborah
; APPLICANT: de Kievit, Teresa
; TITLE OF INVENTION: No. 5994072el Proteins Involved in the Synthesis and Assembly
; FILE REFERENCE: 6580-089
; CURRENT APPLICATION NUMBER: US/08/846,762A
; CURRENT FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 24417
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-08-846-762-1

Query Match      25.7%; Score 306.6; DB 2; Length 24417;
Best Local Similarity 53.9%; Pred. No. 6.9e-92;
Matches 630; Conservative 0; Mismatches 539; Indels 0; Gaps 0;

Qy 22 GCTCTCTCCGACCGGATCTGGGCTGGCGGATCTGTTTGTGCGGATGAACGTCCTCCGCG 81
Db 23576 GCACCGCGTATCCCATCTTGGCTGAACGAAGCTTCAACGCCGATACCCGTCGCGGC 23517
Qy 82 AAAATTAACCTCGGATTTGTCTATAAAGATGAGACGGCAAAACCCCGGTACTGACC 141
Db 23516 AAGATCAACCTGGCGTGGCGGTACTCAACGAGGAGGGCGCATTCCTCGTGTGGGT 23457
Qy 142 AGGCTGAAAAAGGTGAACAGATATCTCTCGAAAAATGAACCAACCAAAAAATTAACCTCGCG 201
Db 23456 GCGCTGAGCTGCCGAGAGGGCCGATCGAGGCCACGCGCGCGCTACCTCGCG 23397
Qy 202 ATTGACGGCATCCCTGAATTTGTGCTGCACTCAGGAATGCTGTTGTTGTAAGGTAGC 261
Db 23396 ATCGAAGGATCGCGCTACGACACGAGGCGGTACAGAAATGCTGTTTCGTTAAGAGTCC 23337
Qy 262 GCCTGATCAATGACAAACGTCGTCACGCGCAGACTCGGGGGGACACTGGCGCACTA 321
Db 23336 GAGCTGTGCGCGCGCGCGTGTGTGTCACGACCCAGGCGGTGCGGCGCGCTC 23277
Qy 322 CGGCTGCTGCCGATTTCTTGGCAAAAAATACCAAGCGTTAAGCGTGTGTGGGTGAGCAAC 381
Db 23276 AAGCTCGCGCGGACTTCTCAAGCGCTGCTGCCCGACGCCACCGTGGCCATCAGCGAC 23217
Qy 382 CCAAGTGGCGCAACCAATAAGAGCGTCTTTAACTCTCAGGTCTGGAAGTTGTTGAATAC 441
Db 23216 CCGAGTGGGAAACACACCGCGCATGTTTGAAGCCGCGCTTCCCGGTGAGAACTAC 23157
Qy 442 GCTTATATGATGGGAAATACACTCTTGAATTCGATGCATGATTAACAGCTGAT 501
Db 23156 CGCTATTACACCGCGCCAGCAATGGCGTGAACCGCGCGCGCTTCTGGAAGACTGAT 23097
Qy 502 GAAGCTCAGGCTGGCGACGTAGTGTCTTCATGCGTGTGCCATAACCCCAACCGGTATC 561
Db 23096 GCGCTCGCGCGCGCTGATGCTGTCGACGCTGCTGCCATATCCGACCGCGGCTC 23037
Qy 562 GACCTACGCTGGAACAAATGGCAACACTCGGCAAACTCTCGTTGAGAAGGTGTTA 621
Db 23036 GATCTCGAGTGGACGACTGGAACAGGTGCTGACGCTGCTCAAGGCCAAGGGCGCAGTG 22977
Qy 622 CGCTGTTGACTTCGCTTACAGGGTTTTTGGCGGTGCTGGAAGATGCTGAAGGA 681
Db 22976 CGTTCTCTGACATCGCTTACCAGGGCTTCGGCAACGGCATCGAGGAAGACGCCGCGC 22917
Qy 682 CTGCGCGCTTTCGCGGCTATGCAATAAGAGTGTGTTGTCAGTTCCTACTCTAAAAAC 741

```

Db	22916	GTGGCGCTGTTGCCCAGCTCGGGCCTGAGCTTCTTCGTGTTTCAGCTCGTTCTCTCCAAATCC	22857
Qy	742	TTTGGGCTGTACAAACAGAGCGTGTGGCGCTTGTACTCTGTGTTGTGCGCAGATGAAACC	801
Db	22856	TTCTCGCTCTATGGCGAAACGCGTCGCGCCCTCTCGATCGTGACCGAAAGCCGCGATGAA	22797
Qy	802	GTTGATCGCGCATTCAGCCMAATGAAGCGGGGATTCGGCTAACTACTCTTAACCCACCA	861
Db	22796	TCGGCCCGGGTGCTGTCACAGGTGAAGCGGGTGAATCCGACCAACTATTTCGAACCCACCG	22737
Qy	862	GCAACAGGGCGTCTCTGTGTGTGCCCCATCTCTGAGCAACGATGCGTACTGTCGATTTGG	921
Db	22736	ACCCAGCGGCCAGCGTCGTCTCTCCGTGCTCAACAGCCCGAACTGCGGSCCCTCTGG	22677
Qy	922	GAAACAAGTGTACTGATATGCGCAGCGTATTTCAGCGTATCGGTCAGTTGTTGTCGAAT	981
Db	22676	GAGCAGGAATCGGGCGAGATGCGCGAACCCGATCCCGCAATCGCCTCGCGATGTCGAG	22617
Qy	982	ACGCTGCAGAAAGAGCGCAACCCGCACTTTCAGCTTTATCATCAAAACAGAAACGGCATG	1041
Db	22616	CAACTGCGACCCACCGCGCCAGCGCGACTTCAGTTTCGTGCGTCGCCACGTGCGCATG	22557
Qy	1042	TTCTCTTCAGTGGCCTGACAAAGAAACAAGTCTGCTCTCGCGGAGAGAGTTGGCGTA	1101
Db	22556	TTCTCTCTATTTCGGGCTGACCGCGCACGAGTCGAGCGCCTGAAGACCGAGTTCGGTATC	22497
Qy	1102	TATGCGGTTGCTTCTGCTCGGTAAATGTGGCCGGGATGACACCAAGATAAATGGCTCCG	1161
Db	22496	TAGCCGCTAGACACGGCCGTATCTCGCTCGCGCGCTGACAAATCGAACCTGGAGACC	22437
Qy	1162	CTGTGCGAAGCGGATTTGGCAGTGTGTA	1190
Db	22436	ATCACAAGGCCATCGTCAGGTCTGTGA	22408

RESULT 11

```

US-09-252-991A-12428
; Sequence 12428: Application US/09252991A
; Patent No. 8531795
;
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12428
; LENGTH: 1242
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12428

```

	Query Match	25.6%	Score 305;	DB 3;	Length 1242;
	Best Local Similarity	53.8%	Pred. No. 3.6e-92;		
	Matches 629;	Conservative 0;	Mismatches 540;	Indels 0;	Gaps 0;
Qy	22	GCTCCTGCCGACCGATTCTCGGCGCTGCGCGATCTGTTTGTGCCGATGAACGTCCTCCGCG	81		
Db	73	GCACCGCGTGATCCCATCTTGGCGCTGAACGAGCCTTCAACGCCGATACCGGTCGCGCG	132		
Qy	82	AAAATTAACTCTCGGATTGGTGTCTATAAGATGAGCGGCGAAACCCCGGTACTGACC	141		
Db	133	AAGATCAACTGGCGTGGCGTGTACTACAACGAGGAGGGCGCATTCCTCGCTGTTCGT	192		
Qy	142	AGCGTGA AAAAGCGTGAAACAGTATCTGCTCGAAATGAACCAACCAAAAATTACTCTCGCG	201		
Db	193	GCCGTGCAGGCTGCGCGAAGCGCCGCA TCGAGGCCCA CGGCGCGCGGGTACTCTCGCG	252		

RESULT 12
US-09-252-991A-12162/c
; Sequence 12162, Application US/09252991A
; Patent No. 6551795

```

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12162
; LENGTH: 1329
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12162

Query Match      25.6%; Score 305; DB 3; Length 1329;
Best Local Similarity 53.8%; Pred. No. 3.7e-92;
Matches 629; Conservative 0; Mismatches 540; Indels 0; Gaps 0;

QY 22 GCTCTGCGCCACCGGATTCTGGCGCTGGCGGATCTGTTTCTGCGGATGAACGTCCGGC 81
DB 1212 GCACCGCGTGATCCCATCTCTTGGCTGAACGAAGCCTTTCAACGCCGATACCCGTCGGGC 1153

QY 82 AAAATTAACCTCGGGATTGGTGTCTATAAGATGAGACGGGCAAAACCCCGGTACTGACC 141
DB 1152 AGATCAACCTGGCGTGGCGGTGTACTCAACAGAGAGGGCGCATTCCTGCTGTGCGT 1093

QY 142 AGCGTGAAGAAGCTGAACAGTATCTGCTCGAAAAATGAACACCAAAAAATTACCTCGGC 201
DB 1092 GCGGTGAGCTGCCAGAGAGCGCGCATCGAGGCCCAACCGCGCGCTACCTCGCG 1033

QY 202 ATTGAGCGGATCCCTGAATTTGGTGTGCTGCACTCAGGAATCTGCTGTTTGTGAAGTAGC 261
DB 1032 ATCGAAGGATCCCGCTACGACACGAGGGGTACAGAAATCTGTTTCGGTAAACGATCC 973

QY 262 GCCTCATCAATCAGAAACGTGTCCGACGGCAGACATCCGGGGGCACTGGCGCACTA 321
DB 972 GAGCTGTGCGCCGCGCGGTGTGTTCAGACCCAGCGCGTTCGGCGGCGCGCGCTC 913

QY 322 CGCGTGTGCTCGGATTTCTGGCAAAAAATPACAGCGTTAAGCGGTGTGGGTGAGCAAC 381
DB 912 AAGCTCGCGCGCATCTCTCAAGCGCTCTGCGCGACGCCACCGTGGCCATCAGGAC 853

QY 382 CCAAGTCGCGCGCAACATAAGAGCGTCTTAACTCTGCAAGTCTGGAAGTTCGTGAATAC 441
DB 852 CCGAGCTGGGAAAAACCAACCGCAGCTGTTCCGAAGCGCGCGCTTCCCGGTGCAGAACTAC 793

QY 442 GCTTATTATGATCGGGAATACACTCTTGACTTCGATGCACTGATTACAGCCTGAAT 501
DB 792 CGCTATTACACCGCGCAGCAATGCGGTGAACCGCGCGCGCTGTGTGGAAGACCTGAAT 733

QY 502 GAAGCTCAGGCTGGCGACGTAGTGTCTTCCATGGGTGTGCGCATACCAACCAACCGGTATC 561
DB 732 GCCCTGCGCGCGCTGCGATCTGTGCTGCAAGCTGTGCGCATATCGACCGCGGTC 673

QY 562 GACCTTACGCTGAAACAATGGCAAAACACTGGCAAACTCTCGTTGAGAAAGCTGGTTA 621
DB 672 GATCTCGAGTGGACACTGGAACAGGTGCTGGAAGTGTCTCAAGGCCAAGGCCACGCTG 613

QY 622 CGCTGTTTGACTTCGCTTACAGGGTGTGCGGTGTGCGGTGTGGAAGAGATGCTGAAGGA 681
DB 612 CGGTTCTCGACATCGCTTACAGGGCTTTCGGCAACGGCATCGAGGAAGACGCGCGCG 553

QY 682 CTGCGGCTTTTCGCGGTATGATATAAGAGCTGATTGTTGCGCAAGTTCCTACTCTAAAAAC 741
DB 552 GTGCGGCTGTTGCCCGACGTGGCGCTGAGGCTTCTTGTTCAGCTCGTTCTCAATTC 493

QY 742 TTTGGCTGTATCAACAGAGCGTGTGGCGCTGTGACTCTGCTGCTGCGACAGTGAAC 801
DB 492 TTTCTGCTCTATGGCGAAGCGGTGCGCGCGCTCTCGATCGTAAACCGAAGCGCGATGAA 433

```

```

QY 802 GTTGATCGCGCATTCAGCCAAATGAAAGCGCGGATTCGCGCTAACTACTCTAAACCCACCA 861
DB 432 TCGGCCCGCGTGTGCTCCAGAGTGAAGCGAGTATCGGACCACTATTGAAACCCACCG 373

QY 862 GCACAGCGCGCTTCTGTTGTGACCATCTCAGCAACGATGGTTACGTGGATTTGG 921
DB 372 ACCACGCGCGCAGCGTCTGCTCTCCGTCTCAACAGACCGCGAATTCGCGGCGCTCTGG 313

QY 922 GAAACAAGAGCTACTGATATGCGCCAGCGTATTTCAGGTATCGGTGTTGCTCAAT 981
DB 312 GAGCAGAACTGGCGAGATGCGGACCGCATTCGCGCATGCGCCTGGCGATGCTCGAG 253

QY 982 ACGCTGAGGAAAAAGCGCAACCGCACTTCAGCTTTATCATCAACAGAAACGGCATG 1041
DB 252 CAACTGCGAGCCCAACGCGCGCAAGCGCACTTCAGCTTCGTCGTCGCCAACGTTGGCATG 193

QY 1042 TTCTCTTCTAGTGGCTGACAAAAGAAACAAAGTCTGCTGCGCGAAGAGTTTGGCGTA 1101
DB 192 TTCTCTTATTCGCGCTGACCGCGCAGCGCTGAGCGCTGAAGACCGAGTTGCGTATC 133

QY 1102 TATGCGGTTGCTTCTGCTGCGTAAATGTGGCGGGATGACACAGATACATGCGCTCCG 1161
DB 132 TAGCGCGTCAGCACCGCGCGTATCTGCGTCCGCGCTGAACAAATCGAACCTGGAGACC 73

QY 1162 CTGTGCGAAGCGATTGTGGCAGTGTGTA 1190
DB 72 ATCACCAGGCCATCTCCAGTCTCTGTA 44

RESULT 13
US-09-489-039A-1301
; Sequence 1301, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1301
; LENGTH: 1338
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1301

Query Match      24.3%; Score 289; DB 3; Length 1338;
Best Local Similarity 53.2%; Pred. No. 9.7e-87;
Matches 635; Conservative 0; Mismatches 555; Indels 3; Gaps 1;

QY 2 TGTGTTGAGAACATTACCGCGCTCTGCGGACCCGATTCGCGCGCTGGCGGCTCTGTTTC 61
DB 146 TGTTCAAAAGTTGAGCTCTACCGCGCGGACCTATTCTTCTCTGATGGAACGCTTCA 205

QY 62 GTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGTTGTTGTTGTTGTTGTTGTTG 121
DB 206 AAGAAGACCCGCGAAGCGACAAAGTCAACTGAGTATCGGGCTGTACTCAACGATGACG 265

QY 122 GCMAAACCCCGGTACTGACCGGTTGAAAAGGCTGAAACAGTATCTGCTCGAAAAATG--- 178
DB 266 GCATTTATCCCGAGCTGCGAGCGGTGGCGGAGCGGAAGCAGCCCTGAAACCGGAGCCG 325

QY 179 AAACACCAAAAAATTACCTCGGATTCGCGCATCCCTGAAATTTGTTGTTGTTGTTGTTGTTG 238
DB 326 ATGCGCGCTCTGCTGTTATCTGCCATGGAAGGGTTGAGCGGCTACCGTTCAGGCGATTCGCG 385

QY 239 AACTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 298
DB 386 CACTGCTGTTGCGCGAGAGCATACCGCGCTTAAGCAAAATCGCAATTCGCTCTATTTCAGA 445

```



```

QY 299 CTCGGGGGCACTGCGGCACTACGCGTGGCTGCCGATTTCTTGGCAAAATACACGCG 358
Db 446 CGGTGCGCGGTTTGGCGCGCTGAAGGTCGGGCGACCTTCTCTCAAAACGTTACTTTCCTG 505
QY 359 TTAAGCGGTGTGGGTGAGCAACCCAAAGCTGGCGCAACATAAGAGCGTCTTTAACTCTG 418
Db 506 AGTCTCATGTCTGGGTACGGATCCGACTTGGGAAACCAACATCGCCATTTTGAAGGG 565
QY 419 CAGGCTCGGAAGTTCGTGAATACGCTTATATGATCGGAAATCAACATCTTGACTCG 478
Db 566 CTGGCTTCGAAGTAAGTACTTACCCCTGTGTTTGTAAAGCCACCAACGCGGTGCGCTTGG 625
QY 479 ATGCACTGATTAACACGCTGAATGAAGCTCAGCTCGCGGACAGTAGTCTGTTCCATGGCT 538
Db 626 AGGACCTGCTGGCGACGCTGCAAAACGCTGCGCGCCGCGATATATGTGCTGCTGCACCCCT 685
QY 539 GCTGCCATAACCAACCGGTATCGACCTACGCTGGAACAAATGGCAAAACACTGGCAAC 598
Db 686 GTGCCACAACCCGACCGGCGCGACTGAGCCAGCCAGTGGGACCGGTTGGTTGAGG 745
QY 599 TCTCCGTTGAGAAGGCTGTTACCGCTGTTTGACTTTCGCTTACAGGGTTTTGCCGCTG 658
Db 746 TATTGAAGCGCTCAGCTGATCCCTTCTCTGACATCGCTTACCAAGGCTTTGGCGGG 805
QY 659 GTCTGGAAGAAGATGCTGAAGGACTGCGGCTTTCGCGGCTATGCAATGAAGAGCTGATTG 718
Db 806 GGTGGAAGGATGCTGCTGCTATTCGCGCCATCGCAGCGCGGGATGCCGATGCTGG 865
QY 719 TTGCCAGTTCCTACTCTAAAACTTTGGCTGTGACAAACGAGCGGTGTGGCGCTGTGACTC 778
Db 866 TCAGCAACTGTTCTTCAAAATTTTCTCGCTGACGCGGAGCGGCTGCGCGGCTCTGTCG 925
QY 779 TGGTTGTCGCCGACAGTGAACCGTTGATCGCGCATTTACGCCAAATGAAGCGCGGATTC 838
Db 926 TGGTCTGTGAAGACAGGAAACCGCGCGCGGTGCTGGGCGAGCTGAAGGCCACCGTGC 985
QY 839 GCCTAACTACTCTAACCAACGACACACGCGCGCTTCTGTGTTGACCATCTGACGA 898
Db 986 GCGGTAACTACTCCAGCCCGCGAGCTTTGGCGCGCAGGTGGTGGCGACGGTCTGAACG 1045
QY 899 ACATGCGTTACGTGGATTTGGAAACAGAGCTGACTGATATGCGCAGCGTATTTCAGC 958
Db 1046 ATCGCGGCTGAAGACCACTGGCAGCGGAGTGACGCCATTCGCGCGCACATCTGTA 1105
QY 959 GTATGCTCAGTTGTTTCTCAATACGCTGCAGGAAAGCGCAACCGCGACTTCAGCT 1018
Db 1106 CCATGCGCCAGCGCTGTCGACGCGCTGCAGCAGGTTGCCCGGCGAGCAAAAGTGGATT 1165
QY 1019 TTATCATCAACAGAACGGCATGTTCTCTTACGTGCGCTGACAAAGAACAAAGTGTGTC 1078
Db 1166 ACCTGCTCAAGCAGCGCGGATGTTCAAGCTATACCGGATTTAGCGCGCGCAGGTGGATC 1225
QY 1079 GTCTGCGGAAGATTTGGCGTATGCGGTTGCTTCTGCTGCGTAAATGTGCGCGGA 1138
Db 1226 GTCTGCGGATGATGTTGGCGTCTACTGATTTGACGCGAGCGCATGTGCTGGCGGGGT 1285
QY 1139 TGACACAGATAACATGGCTCCGCTGTGCGAAGCGATTGTGGCAGTCTGTAA 1191
Db 1286 TAAACTCAGTAAATGTTACAGAGGTAGCAAGCTTTTGGCGCGGTAATGTAA 1338

```

```

RESULT 14
US-09-489-039A-1274/c
; Sequence 1274, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27

```

```

; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1274
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1274

```

```

Query Match 23.4%; Score 278.2; DB 3; Length 1185;
Best Local Similarity 53.4%; Pred. No. 4.1e-83;
Matches 608; Conservative 0; Mismatches 528; Indels 3; Gaps 1;

```

```

QY 2 TGTGTTGAGAACTTACCGCGCTCTGCGACCCGATCTGGGCTTGGCGGATCTGTTC 61
Db 1139 TGTGTTCAAAAAGTTGACGCTACGCGGGGACCTATCTTCTCTGATGGAACGCTTCA 1080
QY 62 GTGCGGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACGG 121
Db 1079 AAGAAGACCCGCGAAGCGACAAAGTCAACCTGAGTATCGGGCTACTACAACGATGACG 1020
QY 122 GCAAAACCCGCTACTGACCGGTGAAAAAGCTGAACAGTATCTGCTCGAAATG--- 178
Db 1019 GCATTTATCCCGAGCTGACGCGGTGGCGAAGCGGAACGACGCTGAACCGCGAGCCGC 960
QY 179 AAACCAACCAAAATTAACCTCGGCTATCGGCAATCCCTGAAATTTGGTTCGCTGCACCTCAGG 238
Db 959 ATGGCGCTCGCTGTATCTGCCAATGGAAGGTTGAGCGGCTACCGTCAAGGCGATTGCCG 900
QY 239 AACTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTTGTCGACGGCACAGA 298
Db 899 CACTGCTGTTTGGCGCAGAGCATACGCGCTTAAAGCAAAATCGCATTTGCTCTATTTCAGA 840
QY 299 CTCGGGGGGCACTGCGGCACTACGCGTGGCTGCGGATTTCTTGGCAAAATACACGCG 358
Db 839 CGGTGCGCGGTTTGGCGCGCTGAAGGTCGCGGCGGCTTCTTCAAAACGTTACTTTCCTG 780
QY 359 TTAAGCGCTGTGGGTGAGCAACCCAAAGCTGGCGCAACCATTAAGAGCGCTCTTTAACTCTG 418
Db 779 AGTCTCATGTCTGGGTACGAGTCTTACCCCTGTTTGTATAAAGCCACCAATCGCCATTTTGAAGGG 720
QY 419 CAGGCTCGGAAGTTCGTGAATACGCTTATATGATCGGAAATCAACATCTTGTGACTTCG 478
Db 719 CTGGCTTCGAAGTAAAGTACTTACCCCTGTTTGTATAAAGCCACCAACGCGGCTGCGCTTG 660
QY 479 ATGCACTGATTAACAGCTGAATGAAGCTCAGCTGCGGACGAGTAGTCTGTTCCATGGCT 538
Db 659 AGGACCTGCTGGCGACGCTGCAAAACGCTGCGCGCCGCGATATTGTGCTGCTGCACCCCTT 600
QY 539 GCTGCCATAACCAACCGGTATCGACCTATCGCTGGAACAAATGGCAAAACACTGGCAAC 598
Db 599 GCTGCCACAACCCGACCGGCGCGACTGACGCGACCCAGTGGGACCGCGTGTGAGG 540
QY 599 TCTCCGTTGAGAAGGCTGTTACCGCTGTTTGACTTTCGCTTACAGGGTTTTGCCGCTG 658
Db 539 TATTGAAGCGCTCAGCTGATCCCTTCTCTGACATCGCTTACAGGGCTTTGGCGGG 480
QY 659 GTCTGGAAGAAGATGCTGAAGGACTGCGGCTTTCGCGGCTATGCATAAAGAGCTGATTG 718
Db 479 GGTGGAAGGATGCTGCTGCTATTCGCGCCATCGCCAGCGCGGGATGCCGATGCTGG 420
QY 719 TTGCCAGTTCCTACTCTAAAACTTTGGCTGTGACAAACGAGCGGTGTGGCGCTTGTATCTC 778
Db 419 TCAGCAACTCGTTCTTCAAAATTTTCTGCTGTACGCGGAGCGGCTCGCGGCTGTGCTCG 360
QY 779 TGGTTGCTGCCGACAGTGAACCGTTGATCGCGCATTTACGCCAAATGAAGCGCGGATTC 838
Db 359 TGGTCTGTGAAGACAGGAAACCGCGCGCGCTGCTGCGGCGAGCTGAAGGCCACCGTGC 300
QY 839 GCCTAACTACTCTAACCAACGACGACGCGGCTTCTGTTGTCACCATCTCTGACGA 898
Db 299 GCGGTAACTACTCCAGCCCGCGAGCTTTGGCGCGCAGGTTGGTGGCGACGCTGCTGAACG 240

```


Qy	899	ACGATGCGTTAGCTGCGATTTGGGNAACAAGAGCTGACTGATATGCCCGCAGCGTATTTCAGC	958
Db	239	ATGCGGGCTGAAAGCCACGTGGCAGCGCGAAGTGGAGCCATGCGCGCACATCTCTGA	180
Qy	959	GTATGCGTCAGTTGTTGCTCAATACGCTGCAGGAAAAAGGCGCAAAACCGCGACTTTCAGCT	1018
Db	179	CCATGCGCCAGCGCTGTCAGCGGCTGCAAGAGTTGCCCGGCGAGCAAGTGGATT	120
Qy	1019	TTATCATCAACAGAACGGCATGTTCTCCTTCAGTGGCCTGCACAAAGAACCAAGTGCTGC	1078
Db	119	ACCTGCTCAAGCAGCGCGGATGTTTCAGCTATACCGGATTTCGCGCGCAGGTGGATC	60
Qy	1079	GTCGCGGGAAGAGTTGGCGTATATGCGGTTGCTTCCTGGTCTGCGTAAATGTGCGCGG	1137
Db	59	GTCGCGCGATGAGTTTGGCGCTACCTGATTGCCAGCGACGATGTGCTGTCGCGGG	1

RESULT 15

US-09-540-236-315

; Sequence 315, Application US/09540236

; Patent No. 6673910

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXE

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2005-001

; CURRENT APPLICATION NUMBER: US/09/540.236

; CURRENT FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 3840

; SEQ ID NO 315

; LENGTH: 1209

; TYPE: DNA

; ORGANISM: M.catarrhalis

US-09-540-236-315

Query Match	23.2%	Score 276.4;	DB 3;	Length 1209;
Best Local Similarity	52.6%	Pred. No. 1.7e-82;		
Matches 626;	Conservative	0;	Mismatches 561;	Indels 3;

[illegible]

Db	493	GATGAGCTGTGCGAATTTTTTCAAGACTTTTAAATGAAATGATGTAGTGCTACTACACCCA	552
Qy	538	TGCTGCCATAACCCAAACCGGTATCGACCCCTACGCTGGAACAATGGCAAAACACTGGCACAA	597
Db	553	TGTTGCCATAACCCCAACTGTGTGGATTTCACCCGTGAACAGTGAATGTACTATTACAA	612
Qy	598	CTCTCCGTTTGAGAAAGCGTGGTTACCGCTGTTTGACTTTCGCTTACCAAGGGTTTTGCGCGT	657
Db	613	ATCGTCAAAAGATAAAAAAACCCTTATTCCAATTTATGGATATTCCTATCAAGGTGTTTGGCGAT	672
Qy	658	GGTCTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGCTATGCATAAAGAGCTGATT	717
Db	673	GATATGATGGTGATGCGTATGCCATTCGTCGTCGCGTGGGCTGGGATGACGGTATTT	732
Qy	718	GTTGCCAGTTCCTACTCTTAAAAACTTTGGCCCTGTACAAACGAGCGGTGTGGCGCTGTGTACT	777
Db	733	GTCTCTAAATCATTTTCAAAAAAATTGTCACTTTATGGTGAGCGGTGTGGCGCGTGTATCT	792
Qy	778	CTGGTTCCTCCGACAGTGAACCGTTGTATCGCGGCATTCAGCCAAATGAAAGCGGCGATT	837
Db	793	GTCTGTGCACCAACCAAGAAGACGATGTGGTACTTGGTCAATTTAAATTTTACGTC	852
Qy	838	CGCGCTAACTACTTAAACCCACGACACACGGCGCTTCTGTTGTGCCACCATCTCTGAGC	897
Db	853	CGTCGCAATTTATCAAGTCCAGCTGCACACGGCAACTATGTTGTTGATCGGTCACTGAAT	912
Qy	898	AACGATCGGTTCGTGGATTTCGGAAACAGAGCTGACTGATATGGCCAGCGTATTTCAG	957
Db	913	GATGAAGCATTTATTTGGGTTATGGATTGTGGAGGCATATGAATATGGTGACCGCATTCGT	972
Qy	958	CGTATGCGTCAGTTGTTTCGTCAATAGCTGCAGGAAAAAGCGCAACACCGCACTTCAGC	1017
Db	973	GATATGCGTCAAAAACTCCAAGAAGCATTAAGCACATAACTGCTCGAGCGTGATTTTAGC	1032
Qy	1018	TTTATCATCAAAACGAAGCGCATGTTCTCCTTCAGTGGCCCTGACAAAGAACAAGTGCTG	1077
Db	1033	TATTTTACCAACACAGCGTGATGTTTAGCTTTACAGGACTAACCAAGAAGCAAGTGATT	1092
Qy	1078	CGTCTGCGCGAAGTTTGGCGTATATGCGGTTTGTCTTCGTGTCGCGTAAATGTGGCCGGG	1137
Db	1093	CGTTTGGCTGAAGNATTTGGCGTATATATGGTTGAAAATGGCGGTATGTGTGTGGCAGGT	1152
Qy	1138	ATGACACCAAGATAACATGGCTCCGCTGTGCGAAGCGGATTTGTGGCAGTGCT	1187
Db	1153	CTGAATAATTTCTAATGTAGATAATGTGTGCGAGTGTCTATGGCTGAAGTGCT	1202

Search completed: March 15, 2006, 03:10:27
Job time : 459 secs

This Page Blank (uspto)

GenCore version 5.1.7.
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 15:26:24 ; Search time 3885 Seconds
(without alignments)
4769.031 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045

Sequence: 1 MFENITAPADPILGLDLF.....VAGTPDNMAPLCEAIVAVL 396

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=exp
-Q=/abss/ABSSWEB_spo01/US10673786/runat_14032006_135611_22627/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05p
-USER=US10673786 @CGN 1 1 8010 @runat_14032006_135611_22627 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hc3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1486	72.7	864	10	CL660432
2	1352	66.1	785	10	CL695845
3	989	48.4	645	6	CA0933007
4	947	46.3	671	10	CL695172
5	914	44.7	1690	4	AY103658
6	909	44.4	1749	8	CX109239
7	868	42.4	1599	4	CNS0A2EC

8	857	41.9	1832	4	AY105250	Zea mays
9	835	40.8	557	6	CF306417	HDAL--03-
10	835	40.8	1486	4	CNS092WQ	BX829537 Arabidops
11	807.5	39.5	1502	4	AY815146	Schistosoma
12	797	39.0	2452	4	CR857622	Pongo pyg
13	795	38.9	468	1	AW221110	EST297579
14	777	38.0	2447	4	CR860094	Pongo pyg
15	759	37.1	1496	4	CR616132	full-leng
16	759	37.1	1509	4	CR626608	full-leng
17	759	37.1	1512	4	CR605482	full-leng
18	759	37.1	1525	4	CR615083	full-leng
19	759	37.1	1540	4	CR598904	full-leng
20	759	37.1	1560	4	CR622320	full-leng
21	759	37.1	1563	4	CR601723	full-leng
22	759	37.1	1719	4	CR606661	full-leng
23	759	37.1	1959	4	CR625122	full-leng
24	759	37.1	1979	4	CR619243	full-leng
25	759	37.1	1979	4	CR620721	full-leng
26	759	37.1	2017	4	CR860603	Pongo pyg
27	754	36.9	1564	4	CNS0EPVU	Tetraodon
28	748	36.6	1517	4	CNS0GNY6	CR725768 Tetraodon
29	746.5	36.5	1510	4	AY223393	Schistosoma
30	745	36.4	1440	4	CNS0EOL3	Tetraodon
31	739.5	36.2	2019	4	CR326063	Pongo pyg
32	736	36.0	1643	4	CNS0A80C	BX820010 Arabidops
33	734	35.9	2050	4	CR861057	Pongo pyg
34	733.5	35.9	977	7	CK259742	EST705820
35	732	35.8	968	7	CK262759	EST708837
36	721	35.3	1704	4	CR625110	full-leng
37	718	35.1	1566	4	CNS0A3OI	Arabidops
38	718	35.1	1762	4	CNS0GRLY	Tetraodon
39	712	34.8	1771	4	CNS0EYV2	CR444084 Tetraodon
40	709	34.7	1034	7	CV734935	Cr12_11.M
41	708	34.6	1358	7	CO635670	Cont1g274
42	706.5	34.5	902	7	CK259741	EST705819
43	705	34.5	902	7	COL13549	GR_EB013
44	702	34.3	1131	7	CV734832	Cr12_11.H
45	697	34.1	1102	10	AY407844	Homo sapi

ALIGNMENTS

CL660432 864 bp DNA linear GSS 09-JUL-2004
PRI0137a.D10 - PRI0137a.B21 (864) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
CL660432
CL660432.1 GI:50145720
GSS.
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 864)
Srinivasan, J., Otto, G.W., Kahlow U., Geisler, R. and Sommer, R.J.
AppaDB: an AcedB database for the nematode satellite organism
Pristionchus pacificus
Nucleic Acids Res. 32 (1), D421-D422 (2004)
14681447
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: 77
Class: fosmid ends.
Location/Qualifiers

FEATURES

```
source
1..864
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
/note="Vector: pEpifos-5 Fosmid vector"

ORIGIN
Alignment Scores:
Pred. No.: 1.36e-160 Length: 864
Score: 1486.00 Matches: 285
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 72.7% Indels: 0
DB: 10 Gaps: 0

US-10-673-786A-2 (1-396) x CL660432 (1-864)
Qy 19 LeuPheArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValTyrlsAsp 38
Db 10 CTGTTTCGCGATGAACGTCCTCCGCAAAATTAACCTCGGATGGTGTCTTAAGAT 69
Qy 39 GluThrGlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrlsLeuGlu 58
Db 70 GAGACGGCAAAACCCCGGTACTGACACGCGTGAAGAGGCTGAACATGATCTGCTCGAA 129
Qy 59 AenGluThrThrLysAenTyrlsLeuGlyIleAspGlyIleProGluPheGlyArgCysThr 78
Db 130 AATGAACCCACCAAAATTAACCTCGGCAATTCACGCGCATCCCTGGAATTTGGTCGTCAC 189
Qy 79 GlnGluLeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAla 98
Db 190 CAGGAACCTGCTTTGGTAAGATGAGCCCTGATCAATGACAAACGTCGTCGACGGCA 249
Qy 99 GlnThrProGlyGlyThrGlyValAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThr 118
Db 250 CAGACTCGGGGGGCACTGCGCACTACGCTGGCTGCGGATTTCTCGCAAAATAAC 309
Qy 119 SerValLysArgValTrrValSerAenProSerTrpProAenHisLysSerValPheAen 138
Db 310 AGCGTTAAGCGTGTGGGTGAGCAACCAACAGCTGGCGCAACCAATAGAGCGCTCTTAAC 369
Qy 139 SerAlaGlyLeuGluValArgGluTyrlsAlaTyrlsAspAlaGluAenHisThrLeuAsp 158
Db 370 TCTGAGGTCTGGAAGTTCTGTAATACCTTTATATGATGCGGAATACACACTCTTGAC 429
Qy 159 PheAspAlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHis 178
Db 430 TTCGATGCACCTGATTAACAGCGCTGAATGAAGCTCAGGCTGGCGAGCTAGTCTGTTCAT 489
Qy 179 GlyCysCysHisAenProThrGlyIleAspProThrLeuGluGlnTrrPheLeuAla 198
Db 490 GGCTGCTGCCATAACCCCAACCGGTATGACCTATGACCTGCGCAACATGCGCAACACTGGCA 549
Qy 199 GlnLeuSerValGluLysGlyTrrLeuProLeuPheAspPheAlaTyrlsGlnGlyPheAla 218
Db 550 CAACCTCTCGTTGAGAAAGCGTGGTTACCGCTGTTTGACTTCGCTTACACGGGTTTGCC 609
Qy 219 ArgGlyLeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeu 238
Db 610 CGTGGTCTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGCTATCATATAAGAGCTG 669
Qy 239 IleValAlaSerSerTyrlsSerLysAenPheGlyLeuTyrlsAenGluArgValGlyAlaCys 258
Db 670 ATTGTTGCGAGTCTCTACTCTAAAAAATTTGGCCCTGTACACGAGCGGTGTTGGCGCTGT 729
Qy 259 ThrLeuValAlaAlaAspSerGluThrValArgAlaPheSerGlnMetLysAlaAla 278
Db 730 ACTCTGGTGTCTGCCACAGTGAACCGTTGATTCGCGCATTCAGGCCAATGAAAGCGCG 789
Qy 279 IleArgAlaAenTyrlsSerAenProAlaHisGlyAlaSerValValAlaThrIleLeu 298
Db 298
```

```
Db 790 ATTGCGCTAACTACTCTAAACCCAGCAGCACGGCGCTTCTGTGTGTCACCATCTG 849
Qy 299 SerAenAspAlaLeu 303
Db 850 AGCAACGATGGTTA 864

RESULT 2
CL695845 785 bp DNA linear GSS 10-JUL-2004
PRI017b.H12.2 - PRI017b.BR (785) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
CL695845
VERSION CL695845.1 GI:50217753
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Rukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE
1 (bases 1 to 785)
AUTHORS Strinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
TITLE AppaDB: an AcedB database for the nematode satellite organism
Pristionchus pacificus
JOURNAL Nucleic Acids Res. 32 (1), D421-D422 (2004)
PUBMED 14681447
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: Fosmid ends.
Location/Qualifiers
1..785
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
/note="Vector: pEpifos-5 Fosmid vector"

FEATURES
source
1..785
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
/note="Vector: pEpifos-5 Fosmid vector"

ORIGIN
Alignment Scores:
Pred. No.: 3.63e-145 Length: 785
Score: 1352.00 Matches: 260
Percent Similarity: 99.6% Conservative: 0
Best Local Similarity: 99.6% Mismatches: 1
Query Match: 66.1% Indels: 0
DB: 10 Gaps: 0

US-10-673-786A-2 (1-396) x CL695845 (1-785)
Qy 128 ProSerTrpProAenHisLysSerValPheAenSerAlaGlyLeuGluValArgGluTyrl 147
Db 3 CCAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGAGGTCTGGAAGTTCTGTGATAC 62
Qy 148 AlaTyrlsAspAlaGluAenHisThrLeuAspPheAspAlaLeuIleAenSerLeuAen 167
Db 63 GCTTATTATGATGCGGAATACACACTCTTGACTTCGATTCGACTGATTAACAGCTGAAT 122
Qy 168 GluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCysHisAenProThrGlyIle 187
Db 123 GAAGCTCAGGCTGGCGAGTAGTGTCTTCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 182
Qy 188 AspProThrLeuGluGlnTrrPheLeuAlaGlnLeuSerValGluLysGlyTrrLeu 207
Db 183 GACCTTACGCTGGCAACATGGCAAACTGGCACAACTCTCCGTTGAGAAAGGCTGGTTA 242
```

QY 208 ProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGlyLeuGluGluAspAlaGluGly 227
 |||||
 Db 243 CGCGCTGTTGACTTCGCTTACCAGGGTTTGGCCGGTGTCTGAAGAAGATGCTGAAGA 302
 |||||
 QY 228 LeuAtqAlaPheAlaMetHisLysGluLeuLeuValAlaSerSerTyrSerLysAsn 247
 |||||
 Db 303 CTGCGCGCTTTGCGGCTATGCATAAAGAGCTGATTGTTGCCAGTTCCTACTCTAAAAAC 362
 |||||
 QY 248 PheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeuValAlaAlaAspSerGluThr 267
 |||||
 Db 363 TTGGGCTGTACACAGAGCGTGTGGCGCTTGTACTCTGGTGTCTGCCACAGTGAACCC 422
 |||||
 QY 268 ValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAlaAsnTyrSerAsnProPro 287
 |||||
 Db 423 GTTGATCGCGCATTCAGCCAAATGAAAGCGCGGATTCGGCTAACTACTCTAACCCACCA 482
 |||||
 QY 288 AlaHisGlyValSerValAlaThrIleLeuSerAsnAspAlaLeuArgAlaIleTTP 307
 |||||
 Db 483 GCACAGCGCGCTTCTGTTGTCACCATCTGAGCAACGATGCGTTAGTGGATTTGG 542
 |||||
 QY 308 GluGlnGluLeuThrAspMetArgGlnArgIleGlnArgMetArgGlnLeuPheValAsn 327
 |||||
 Db 543 GAACAGAGCTGACTGATATGCCACGATATTCAGCGTATGCGTCAGTTGTTCGTCAT 602
 |||||
 QY 328 ThrLeuGlnGluLysGlyAlaAsnArgAspPheIleIleLysGlnAsnGlyMet 347
 |||||
 Db 603 ACGTGCAGGAAAGGCGCAACCGCGACTTCAGCTTTATCATCAACAGAGGCGATG 662
 |||||
 QY 348 PheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeuArgGluGluPheGlyVal 367
 |||||
 Db 663 TTCTCCTTCAGTGGCGCTGACAAAGAACAAAGTGTGCGTCTGCGCGAAGATTTGGCGTA 722
 |||||
 QY 368 TyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThrProAspAsnMetAlaPro 387
 |||||
 Db 723 TATCGGTTGCTTCTGTCGCTAAATGTGGCGGATGACACAGATAACATGTTCCG 782
 |||||
 QY 388 Leu 388
 |||||
 Db 783 CTG 785

RESULT 3
 CA093007
 LOCUS
 DEFINITION 645 bp mRNA linear EST 23-SRP-2003
 3', mRNA sequence.
 CA093007
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

CA093007
 SCCCL1002F08.b Cl1 Saccharum officinarum cDNA clone SCCCL1002F08
 3', mRNA sequence.
 CA093007.1 GI:34946314
 EST.
 Saccharum officinarum
 Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Saccharum; Saccharum officinarum
 complex.

REFERENCE
 1 (bases 1 to 645)
 Vettore A.L., da Silva F.R., Kemper E.L. and Arruda, P.
 The libraries that made SUCEST
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parruda@unicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bccc.org.br
 Plate: 002 row: F column: 08
 Seq primer: M13/Forward primer.
 Location/Qualifiers
 1. .645

FEATURES
 source

RESULT 4
 CL695172
 LOCUS

CL695172 671 bp DNA linear GSS 10-JUL-2004

/organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCCCL1002F08"
 /lab_host="XL1Blue MRF"
 /clone_lib="Cl1"
 /note="Organ: Pool of sugarcane calli submitted to low
 (40C) and high (37 C) temperature stress; Vector:
 pBluscript; Site 1: EcoRI; Site 2: XhoI; An
 unidirectional cDNA library generated from [Pool of
 sugarcane calli submitted to low (40C) and high (37 C)
 temperature stress]. cDNA was prepared from polyA+ mRNA
 using ZAP - cDNA Synthesis kit (Stratagene). The
 double-strand cDNAs were fractionated in a sphaerose CL-2B
 40cm-columns and fragments sizing between 0.8 and 1.5 Kb
 were directionally cloned into the vector. Details of
 each source of RNA and library construction can be
 obtained at http://sucet.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:
 Pred. No.: 2,47e-103 Length: 645
 Score: 989.00 Matches: 186
 Percent Similarity: 97.9% Conservatives: 1
 Best Local Similarity: 97.4% Mismatches: 4
 Query Match: 48.4% Indels: 0
 DB: 6 Gaps: 0
 US-10-673-786A-2 (1-396) x CA093007 (1-645)
 QY 31 LeuGlyIleGlyValTyrLysAspGluThrGlyLysThrProValLeuThrSerValLys 50
 Db 22 CTCGGATGTTGTTCTATAAAGATGAGACGGGCAAAACCCCGTACTGACCGCGTGAA 81
 QY 51 LysAlaGluGlnTyrLeuLeuGluAsnGluThrThrLysAsnTyrLeuGlyIleAspGly 70
 Db 82 AAGCGTGAACAGTATCTGCTCGAAAATGAAACACCAAAATTTACCTCGGCATTTACGGC 141
 QY 71 IleProGluPheGlyArgCysThrGlnGluLeuLeuPheGlyLysGlySerAlaLeuIle 90
 Db 142 ATCCCTGATTTGGTTCGCTGCTCAGGACCTGCTGTTTGGTAAGGTAGCGCCCTGATC 201
 QY 91 AsnAspLysArgAlaArgThrAlaGlnThrProGlyGlyThrGlyAlaLeuArgValAla 110
 Db 202 AATGACAAACGTGCTCGACGGCAGACTCCCGGGGGCCTGCGGCACCTAGCGGTGGCT 261
 QY 111 AlaAspPheLeuAlaLysAsnThrSerValLysArgValTyrValSerAsnProSerTTP 130
 Db 262 GCCGATTTCTGGCAAAATACCGAGCTTAAGCGTGTGGTGAGCAACCCCAAGCTGG 321
 QY 131 ProAsnHisLysSerValPheAsnSerAlaGlyLeuGluValArgGluTyrAlaTyrTyr 150
 Db 322 CCGAACCATTAAGAGCGTCTTTAACTCTGACAGTCTGGAAGTTCGTGAATACCGTTATTAT 381
 QY 151 AspAlaGluAsnHisThrLeuAspPheAspAlaLeuIleAsnSerLeuAsnGluAlaGln 170
 Db 382 GATCGGAAATACACACTCTTGACTTCGATGCTGATTAACAGCCTGAATGAAGCTCAT 441
 QY 171 AlaGlyAspValValLeuPheHisGlyCysCysHisAsnProThrGlyLysAspProThr 190
 Db 442 GCTGGCGACGTAGTGTCTGTTCCATGGCTGCTGCCATAACCCCAACCGGTATACGACCC 501
 QY 191 LeuGluGlnTTPGlnThrLeuAlaGlnLeuSerValGluLysGlyTyrLeuProLeuPhe 210
 Db 502 CTGGAAACATGGCAACACTGCGCACTCTCCGTTGAGAAAGGCTGTGTACCGCTGTTT 561
 QY 211 AspPheAlaTyrGlnGlyPheAlaArgGlyLeu 221
 Db 562 GACTTCGCTTACCAAGGTTGCCCGCGGTGAGTT 594

DEFINITION PR10166a_G11.2 - PR10166a.BR (671) Mixed stage fosmid library of *P. pacificus* var. *Californica* Pristionchus pacificus genomic, genomic survey sequence.

ACCESSION CL695172

VERSION CL695172.1 GI:502117080

KEYWORDS GSS.

SOURCE Pristionchus pacificus

ORGANISM Pristionchus pacificus

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.

AUTHORS Srinivasan, J., Otto, G.W., Kahlow, U., Geisler, R., and Sommer, R.J.

TITLE AppaDB: an AcedB database for the nematode satellite organism

JOURNAL Pristionchus pacificus

PUBMED 14681447

COMMENT Nucleic Acids Res. 32 (1), D421-D422 (2004)

Contact: Sommer RJ

Evolutionary Biology

Max-Planck-Institute for Developmental Biology

Spemannstr. 37-39, Tuebingen D-72076, Germany

Tel.: 00497071601371

Fax: 00497071601498

Email: ralf.sommer@uebingen.mpg.de

This library was generated at Caltech, Pasadena, USA and end sequenced at Vancouver, Canada.

Seq primer: T7

Class: fosmid ends.

FEATURES

source

1..671

Location/Qualifiers

/organism="Pristionchus pacificus"

/mol_type="genomic DNA"

/strains="Californica"

/db_xref="taxon:54126"

/clone_lib="Mixed stage fosmid library of *P. pacificus* var. *Californica*"

/note="Vector: pBpifos-5 Fosmid vector"

ORIGIN

Alignment Scores:

Pred. No.: 1,88e-98 Length: 671

Score: 947.00 Matches: 190

Percent Similarity: 97.9% Conservative: 0

Best Local Similarity: 97.9% Mismatches: 3

Query Match: 46.3% Indels: 1

DB: 10 Gaps: 0

US-10-673-786A-2 (1-396) x CL695172 (1-671)

Qy 204 LysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGlyLeuGluGlu 223

Db 1 AAAGGCTGGTTACCGCTGTTGACTTCGCTTACCAGGGTTTTGCCCGTGGTCTGGAGAA 60

Qy 224 AspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuLeuValalaSerSer 243

Db 61 GATGCTGAAGGACTGCGGCGCTTCGCGCTATGCATAAAGAGCTGATTGTCGCCAGTTC 120

Qy 244 TyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeuValalaAla 263

Db 121 TACTCTAAAAACTTTGGCGCTGTACACGAGGCTGTGGCGCTGTGACTCTCGTTGCTGCC 180

Qy 264 AspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAlaAsnTyr 283

Db 181 GACAGTGAACCGTTGATCGCGCATTCAGCAAAATGAAGCGGCGATTGCGCGCTAACTAC 240

Qy 284 SerAsnProProAlaHisGlyAlaSerValValalaThrIleLeuSerAsnAspAlaLeu 303

Db 241 TCTAACCCACAGACACCGGCGCTCTCTGTTGTCACCATTCCTGAGCAACGATCGGTTA 300

Qy 304 ArgAlaIleTpGlu-GlnGluLeuThrAspMetArgGlnArgIleGlnArgMetArgG1 323

Db 301 CGGCCCGATTGGGAAACACAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGTATGCGTCA 360

Qy 323 nLeuPheValaAsnThrLeuGlnGluLysGlyAlaAsnArgAspPhePheIleLeLy 343

Db 361 GTTCTTCGTAATACGCTGCAGGAAAAGCGCAAAACCGGACTTCAGCTTTATCATCA 420

Qy 343 sGlnAsnGlyMetPheSerGlyLeuThrLysGluGlnValLeuArgLeuArgG1 363

Db 421 ACAGAAGCGCATGTTCTCTTCAGTGGCTGACAAAAGAACAACTGCTGCTGCGGA 480

Qy 363 uGluPheGlyValTyrAlaValalaSerGlyArgValaAsnValalaGlyMetThrProAs 383

Db 481 AGAGTTTGGCGTATATCGGTTGCTTCTGTCGCTAAATGTCGCGGATGACACCAGA 540

Qy 383 pAsnMetAlaProLeuCysGluAlaIleValalaValLeu 396

Db 541 TAACATGTTCCGCTGTGCGAAGCGATTGTGGCAGTGTG 580

RESULT 5

AV103658 1690 bp mRNA linear HTC 18-FEB-2005

LOCUS AV103658

DEFINITION Zea mays PC0118382 mRNA sequence.

ACCESSION AV103658

VERSION AV103658.1 GI:21206736

KEYWORDS HTC.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

AUTHORS 1 (bases 1 to 1690)

Gardiner, J., Schroeder, S., Polacco, M.L., Sanchez-Villeda, H., Fang, Z., Morgante, M., Landewe, T., Fengler, K., Usec, F., Hanafey, M., Tingey, S., Chou, H., Wing, R., Soderlund, C., and Coe, E.H. Jr.

TITLE Anchoring 9,371 maize expressed sequence tagged unigenes to the bacterial artificial chromosome contig map by two-dimensional overgo hybridization

JOURNAL Plant Physiol. 134 (4), 1317-1326 (2004)

PUBMED 15020742

REFERENCE 2 (bases 1 to 1690)

Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M., and Tingey, S.V.

AUTHORS Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes

TITLE Unpublished (2002)

JOURNAL 3 (bases 1 to 1690)

REFERENCE 3 (bases 1 to 1690)

AUTHORS Coe, E.H.

TITLE Direct Submission

JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES

Location/Qualifiers

1..1690

/organism="Zea mays"

/mol_type="mRNA"

/db_xref="taxon:4577"

/db_xref="MaizeGDB:636895"

/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"

/notes="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Alignment Scores: 5e-94 Length: 1690

Pred. No.:

Db	1203	ATTAACATGAGGCAACAAC	TATTTATGCGCTGAAATCAGAGGAACCCCTGGTGATTG	1262
Qy	339	SerPheIleIleLyysGlnAsnGlyMetPheSerPheSerGlyLeuThrLyysGluGlnVal	358	
Db	1263	AGCCATATCATTAAGCAAAATGGGATGTTTACTTTCACCTGGCTGATACCGNACAATC	1322	
Qy	359	LeuArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAla	378	
Db	1323	GCATTCATGAGGCAAGAAATACCAACATTTATATGACATCTGTATGGGAGGATCAGCATGGCC	1382	
Qy	379	GlyMetThrProAspAsnMetAlaProLeuCyysGluAlaIleValAlaValLeu	396	
Db	1383	GGTTTGAGCATGAGGACTGTGCCCCATCTTCGAGATGCCATACACGCTCAGTT	1436	

RESULT 6	CX109239	LOCUS	DEFINITION	1749 bp	mRNA	linear	EST 03-JUN-2005
RECOM0798	A normalized whole-life-cycle cDNA library of rice	Oryza sativa (indica cultivar-group)	cDNA clone BI130A13, E1078H06, BI100J15, BI107E19, BI131F09, BI075E21, BI07 5', mRNA sequence.				
ACCESSION	CX109239	VERSION	CX109239.1	GI:66922391			
KEYWORDS	EST.	SOURCE	Oryza sativa (indica cultivar-group)				
ORGANISM	Oryza sativa (indica cultivar-group)						
REFERENCE	Zhang,J., Feng,Q., Jin,C., Qiu,D., Zhang,L., Xie,K., Yuan,D., Han,B., Zhang,Q. and Wang,S.						
AUTHORS	Features of ESTs from a normalized cDNA library of the elite indica rice cultivar Minghui 63						
TITLE	Plant J. 42 (5), 772-780 (2005)						
JOURNAL	1591889						
PUBLISHED	Contact: Wang S						
COMMENT	National Key Laboratory of Crop Genetic Improvement						
	Hauzhong Agricultural University						
	Wuhan 430070, China						
	Tel: 86-27-87282044						
	Fax: 86-27-87287092						
	Email: shipingwang@hotmail.com						
	Seq primer: T7.						
FEATURES	Location/Qualifiers						
source	1..1749						
	/organism="Oryza sativa (indica cultivar-group)"						
	/mol_type="mRNA"						
	/strain="indica"						
	/cultivar="Minghui 63"						
	/db_xref="taxon:39946"						
	/clone="BI130A13, E1078H06, BI100J15, BI107E19, BI131F09, BI075E21, BI07"						
	/tissue_type="whole plant"						
	/dev_stage="whole-life-cycle"						
	/lab_host="E. coli DH10B"						
	/clone_lib="A normalized whole-life-cycle cDNA library of rice"						
	/note="Vector: pSPORN1; Site 1: SalI; Site 2: NotI; The library is constructed based on the strategy of saturation hybridization with genomic DNA using rice cultivar Minghui 63. This library consists of cDNA from 15 directionally cloned cDNA libraries constructed with different tissues from 9 developmental stages."						

ORIGIN	Alignment Scores:	Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:
1..99e-93	Length:	1749				
309.00	Matches:	179				
62.3%	Conservative:	69				
45.0%	Mismatches:	148				
44.4%	Indels:	2				

DB:	8	Gaps:	2
US-10-673-786A-2 (1-396) x CX109239 (1-1749)			
Qy	1	MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe	20
Db	281	GTCTTCGCTGGCTCGCGAGCGCGGAGGAGCCCATCTCCGCGGTGACGGTCCGCTG	340
Qy	21	ArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValTyLysAspGluThr	40
Db	341	AACAAGGATCCAGCCCGTGAAGTCAATCTCGGCGTCCGCGCTACCGGACGAGAA	400
Qy	41	GlyLysThrProValLeuThrSerValLysIleAlaGluGlnTyLysLeuGluAenGlu	60
Db	401	GGGAAGCCCTGGTCTGAATGTGTGAGCGCGCTGAGCAGATCTCTGATCAACAACCG	460
Qy	61	Thr---ThrLysAenTyLysGlyIleAenGlyIleProGluPheGlyArgCysThrGln	79
Db	461	TCACGTGTAGGAGTATTTCGCGATCACTGACTGGCGGATTTCAATAAGCTGAGTCT	520
Qy	80	GluLeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGln	99
Db	521	AAGCTTATTTTGGTCTGACAGTCTGCCATTCAAGAGAATAGGTGGCTACAGTTTCAG	580
Qy	100	ThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSer	119
Db	581	TGCTTGTGAGAACTGGTTCTTTAAGGTGGAGGTGAATTTCTTTCGAAGGCATTTATCAT	640
Qy	120	ValLysArgValTrrValSerAenProSerTrpProAenHisLysSerValPheAenSer	139
Db	641	GNACGACTATATACATCCACAGCAACCTGGGGGAATCACCCCAAGTGTTCACTTTA	700
Qy	140	AlaGlyLeuGluValArgGluTyTrAlaTyTrAspAlaGluAenHisThrLeuAspPhe	159
Db	701	GCTGCGCTGACTGTAGGATTACCGTACTATGATCTCGCAACCGGTGGACTGATTC	760
Qy	160	AspAlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGly	179
Db	761	CAAGGCGTGTAGAAGATCTCGTTCAGTCTCCTTCAGGTGCAATTTGTTACTGCTTCATGCT	820
Qy	180	CysCysHisAenProThrGlyIleAspProThrLeuGluGlnTyTrGlnThrLeuAlaGln	199
Db	821	TGTGCCCAACCCCTACTGGAGTAGACCACTTTGGACCATGGTGGGAACACATCAGGCG	880
Qy	200	LeuSerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyTrGlnGlyPheAlaArg	219
Db	881	TTGATGAGATCAAAAGCATTCGTGCATCTTTGATACGCTTATCAGGGATTTCGAAGT	940
Qy	220	Gly---LeuGluGluAspAlaGlyLeuArgAlaPheAlaAlaMetHisLysGluLeu	238
Db	941	GGAAGTCTTGACCAAGATGCTCAATCAGTGGCGCATGTTTGTGTGATGGTGGTGAATTG	1000
Qy	239	IleValAlaSerSerTySerLysAenPheGlyLeuTyTrAenGluAenTyValGlyAlaCys	258
Db	1001	CTCAGGCTCAGAGTACGTAGCAATGAACATGGATGTATGAGAGCGGTGCGGTCTTTA	1060
Qy	259	ThrLeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAla	278
Db	1061	AGCATAGTTTGGGAAGTGTGATGATGCTAGCTGTACGGTTGAAAGTCAACTTAAGCTTGA	1120
Qy	279	IleArgAlaAenTySerAenProAlaHisGlyAlaSerValValAlaThrIleLeu	298
Db	1121	ATTAGGCTATGTATTCAAAACCTCTATTATGTTGTCATCTATCGTGGCTACCATACTT	1180
Qy	299	SerAenAspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIle	318
Db	1181	AAGGACAGTGCATTTCAATGAATGGAGCTGTGGAGCTGAAGGGCATCGCTGATAGGATT	1240
Qy	319	GlnArgMetArgGlnLeuPheValAenThrLeuGlnGluLysGlyAlaAenArgAspPhe	338
Db	1241	ATTAGCATGAGGCAACACACTTTTGTGCTCTGAAAACGAGAGAACTCCTGGAGACTGG	1300
Qy	339	SerPheIleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnVal	358

1301 AGTCACATCAATTAGCAGATCGGAATGTTCACTTTCACTGGGCTCAACAGTGAAGTG 1360

359 LeuArgLeuArgGluGluPheGlyValTyAlaAlaSerGlyArgValAenValala 378

1361 GCITTCATGAGGAGGAATACCATTTACATGACATCTGATGGAGGATCAGTATGCT 1420

379 GlyMetThrProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396

1421 GGTTTGAGCGGAGGACCATTCGCGCATCTTTCAGATGCAATCCATGCCGCGATC 1474

CNS0A2EC 1599 bp mRNA linear HTC 06-FEB-2004

Arabidopsis thaliana Full-length cDNA Complete sequence from clone

GSITPGH60ZA09 of Hormone Treated Callus of strain col-0 of

Arabidopsis thaliana (thale cress).

EX828177.1 GI:42459989

HTC; GSLT cDNA.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons: core eudicotyledons;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1599)

Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,

Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,

Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:

A Combined Approach to Evaluate and Improve Arabidopsis Genome

Annotation

Unpublished

2 (bases 1 to 1599)

Genoscope.

Direct Submission

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out

full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli

V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis

genome released by MIPS (Munich Information center for Protein

Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque_EP/Full

length

http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES

source

Location/Qualifiers

1..1599

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/clone="GSITPGH60ZA09"

/tissue_type="Hormone Treated Callus"

/ecotype="Col-0"

/plasmid="pCMVSPORT_6"

complement(1..1599)

/gene="At4g31990"

ORIGIN

Alignment Scores:

Pred. No.: 9.54e-89 Length: 1599

Score: 868.00 Matches: 176

Percent Similarity: 62.2% Conservative: 69

Best Local Similarity: 44.7% Mismatches: 147

Query Match: 42.4% Indels: 2

DB: 4 Gaps: 2

US-10-673-786A-2 (1-396) x CNS0A2EC (1-1599)

Qy	2	PheGluAsnIleThrAlaAlaProAlaAppProIleuGlyLeuAlaAppLeuPheArg	21
Db	218	TTTGAGGGTATAAACAATATGGCTCCACACAGACCCTATCTTTGGAGTCAATTAGCAATTCAAA	277
Qy	22	AlaaspGluArgProGlyLyseIleAsnLeuGlyIleGlyValTyrLyseAspGluThrGly	41
Db	278	GCTGACACTAACGGGATGAACACTCAATCTTGGTGTGGTCTTATCGTACTGAGGAATCTC	337
Qy	42	LysThrProValLeuThrSerValLyseAlaGluGlnTyrLeuLeuGluAsnGluThr	61
Db	338	CAGCCTTATGTCTTAATGTGTATAAAGCGCGAGAATTGTATGTTGGAGAGAGAGAT	397
Qy	62	ThrLyseAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGluLeu	81
Db	398	AACAAAGAGTATCTTCCAATTGAGGGGTGGCAGCAITCAACAAGGCTACTCTGCTGAGTTG	457
Qy	82	LeuPheGlyLyseGlySerAlaLeuIleAsnAspLyseArgAlaArgThrAlaGlnThrPro	101
Db	458	CTATTTTGGAGCTGTGTCATCTCTGTTATTAAAGAACAAAGAGTAGCAACAATTCAGGGTCTT	517
Qy	102	GlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLyseAsnThrSerValLyse	121
Db	518	TCGGGAACAGGTTCACTCGCATTAGCAGCGCTCTTATAGAGCGTTATTTCCTCGGAGCA	577
Qy	122	ArgValTrpValSerAsnProSerTrpProAsnHisLyseSerValPheAsnSerAlaGly	141
Db	578	AAAGTTGTGATCTCATCACCACCTGGGGTAATCACAAGAATATCTTCAATGATGCCAAA	637
Qy	142	LeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAspAla	161
Db	638	GTTCCTGGTGGTCCGAATACCGCTACTATGATCCAAAAACAATTTGGTATTGAGGGA	697
Qy	162	LeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCys	181
Db	698	ATGATACAGATATAAAGGAAGCTCCAGAGGATCCTTTCATCTTGTCTTACGGAATGTCT	757
Qy	182	HisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeuSer	201
Db	758	CACAAACCACTGGAAATTGACCCCAACACAGACAGTGGGTAAAAATTGCTGATGTCAAT	817
Qy	202	ValGluLyseGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly---	220
Db	818	CAGGAAGAAGCAATATCCCATTTTCGATGTTGCATACCAAGGCTTTGCTAGTGAAGC	877
Qy	221	LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLyseGluLeuIleVal	240
Db	878	CTTGATGAAGATGCAGCATCTGTGAGATATTATGCTGACGGGGAATGGAAGTTTGTGT	937
Qy	241	AlaSerSerTyrSerLyseAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu	260
Db	938	GCTCAGTCATATAGTAAAAATTAGTTTGTATGCCAAGAAGATTGGGCAATCAATGTC	997
Qy	261	ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLyseAlaIleArg	280
Db	998	GTGTGCTCCTCAGCTGATGCTGTCTACAAGGCTCAAGGCGAGTTGAAGAGATTGCTCGG	1057
Qy	281	AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn	300
Db	1058	CCTATGTACTCGAATCCACCATTTTCATGGGCGAGAAATCGTGGCCAAATGATGGGTGAT	1117
Qy	301	AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg	320
Db	1118	GTAACATATGTTGATGTAATGAAAGCAGAGATGGAATGATGGCAGGAAGAAATAAGACG	1177
Qy	321	MetArgGlnLeuPheValAsnThrLeuGlnGluLyseGly---AlaAsnArgAspPheSer	339
Db	1178	GTTAGACAAGCTGTATGATAGCTCTGTTTCAAAAGCAAGACGCGGGAAGGACTGTGTC	1237
Qy	340	PheIleIleLyseGlnAsnGlyMetPheSerPheSerGlyLeuThrLyseGluGlnValLeu	359
Db	1238	TTCAATCTGAAGCAAAATGGCATGTTCTCTTTCACCGGCTTAAACAAGCTCAGAGCGAT	1297

Oy	360	ArgLeu	argGlu	cLu	pHe	GlyValTyrAlaValAlaSerCly	argValAsnValAlaGly	379
		:::	:	:	:	:	:	
Db	1298	AACAT	CAGCAAAATGGCATGTGTATGACTAAAGACGGGAGGATATCATTTGGCCCGGA	1357				
Oy	380	MetThrPro	aspAsnMetal	apLoLeu	CysGlu	AlaIleVal	393	
		:::	:	:	:	:	:	
Db	1358	TTCATCTCTGCCAAATGCAGTAGTATCTTGCTGATGCGCATCATC	1399					

RESULT 8

[illegible]

REFERENCES

GARDINER, J., SCHROEDER, S., POLACCO, M. L., SANCHEZ-VILLEDA, H.,
 FANG, Z., MORGANTE, M., LANDEWE, T., FENGLE, K., USECHE, F.,
 HANAFAY, M., TINGEY, S., CHOU, H., WING, R., SODERLUND, C. and COE, B. H.
 JR.

TITLE

TITLE Anchoring 9,371 maize expressed sequence tagged unigenes to the bacterial artificial chromosome contig map by two-dimensional overgo hybridization

JOURNAL Plant Physiol. 134 (4), 1317-1326 (2004)

PUBMED 15020742

REFERENCES

AUTHORS
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.

TITLE

Overgo Probes
Unpublished (2002)
JOURNAL
REFERENCE 3 (bases 1 to 1832)
AUTHORS Coe, E. H.

TITLE

JOURNAL COMMENT

Submitted: (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org: ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES

source

ORIGIN

Alignment Scores:		
Pred. No.:	2.19e-87	Length:
Score:	857.00	Matches:
Percent Similarity:	62.5%	Conservative:
Best Local Similarity:	44.1%	Mismatches:
Query Match:	41.9%	Indels:
DB:	4	Gaps:
		3
		2
		1832


```

Db      194  CTTGTTGACTTCGCTTACCAGGGTTTGGCCGGTCTGGAGAGATGCTGAAGGACTG 253
Qy      229  ArgAlaPheAlaAlaMethHisLysGluLeuLeuValAlaSerSerTyrSerLysAsnPhe 248
Db      254  CGCGCTTCGCGGTATGATTAAGAGAGCTGATTGTTCAGGTTCTCTACTCTAAAAACTTT 313
Qy      249  GlyLeuTyrAsnGluArgValGlyAlaCysThrLeuValAlaAlaAspSerGluThrVal 268
Db      314  CGCCTGTACAAACAGCGTGTGGCGCTTGTACTCTGTTGCTGCGACAGTGAACCGTT 373
Qy      269  AspArgAlaPheSerGlnMetLysAlaAlaLeuArgAlaAsnTyrSerAsnProProAla 288
Db      374  GATCGCGCATTCAGCCAAATGAAGCGCGGATTCGCGCTTAACCTACTCTAACCCACGACA 433
Qy      289  HisGlyAlaSerValValAlaThrLeuSerAsnAlaAlaLeuArgAlaLeuTrpGlu 308
Db      434  CAGCGCGCTTCCTGTTGTTGCCACCATCTTCAGCAAGATGCGTTAGCTGGATTGGGNA 493
Qy      309  GlnGluLeuThrAspMetArgGlnArgGlnMetArgGlnLeuPheValAsnThr 328
Db      494  CAAGAGTGAATGATGATGCGCCAGCGTATTTCAGCGTATGCGTCAGTTGTTGCTCAATACG 553
Qy      329  Leu 329
Db      554  CTG 556

CNS09ZWQ      1486 bp      mRNA      linear      HTC 06-FEB-2004
LOCUS      Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION      GSLTFFB19ZF03 of Flowers and buds of strain col-0 of Arabidopsis
                thaliana (thale cress).
ACCESSION      BX829537
VERSION        BX829537.1 GI:42456349
KEYWORDS       HTC; GSLT cDNA.
SOURCE         Arabidopsis thaliana (thale cress)
ORGANISM       Arabidopsis thaliana
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
                rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE      1 (bases 1 to 1486).
AUTHORS        Castell, V., Aury, J.M., Jallou, O., Wincker, P., Clepet, C.,
                Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
                Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
TITLE          Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
                A Combined Approach to Evaluate and Improve Arabidopsis Genome
                Annotation
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 1486)
AUTHORS        Genoscope.
JOURNAL        Genoscope.
TITLE          Direct Submission
COMMENT        Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
                BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr
                - Web : www.genoscope.cns.fr)
                The sequences are based on single pass reads.
                Life Technologies (a division of Invitrogen) members carried out
                full-length libraries construction : Temple G.
                Genoscope members carried out sequencing and annotation : Castell
                V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C.,
                Schachter V., Weissenbach J., Salanoubat M.
                URV INRA : Clepet C., Caboche M.
                Annotation is based on the June 2003 version of the Arabidopsis
                genome released by MIPS (Munich Information center for Protein
                Sequences). 5 prime and 3 prime are assembled with Phrap.
                http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full-length
                http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
                Location/Qualifiers
                1..1486
                /organism="Arabidopsis thaliana"
                /mol_type="mRNA"
                /db_xref="taxon:3702"
FEATURES       source

```

```

/clone="GSLTFFB19ZF03"
/tissue_type="Flowers and buds"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1..1486
/gene="At5g19550"

gene
ORIGIN
Alignment Scores:
Pred. No.:      5,57e-85      Length:      1486
Score:          835.00      Matches:     176
Percent Similarity: 60.7%      Conservative: 68
Best Local Similarity: 43.8%      Mismatches:  152
Query Match:    40.8%      Indels:       6
DB:              4          Gaps:         2

US-10-673-786A-2 (1-396) x CNS09ZWQ (1-1486)
Qy      1  MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db      99  GTCTTCTCTAAGCTTGTCTCGTCTCCGAAGATCCTATTCTCGGTGAACCTGTTGCTTAC 158
Qy      21  ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db      159  AACAAATGATCCTAGTCTCTGTTAAGATCAATTTGGGTGTCGTCCTACCGAACTGAGAA 218
Qy      41  GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db      219  GGGAAAGCCTCTTGTCTTGTGATGTGTGTAGGAGAGAGCAACACAGTAGTAGAAGATCCG 278
Qy      61  Thr---ThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGln 79
Db      279  TCTCGGGTCAAGGAATACATCCCATTTGTTGTTATTTCCGATTTTAAACAAGTTAAAGCGCC 338
Qy      80  GluLeuLeuPheGlyLysGlySerAlaLeuLeuAspLysArgAlaArgThrAlaGln 99
Db      339  AAGTCATCTTAGGTGCTGACAGTCTCTCGCATTAACAAGATAGAGTTACTACAGTCCAC 398
Qy      100  ThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSer 119
Db      399  TGTCTGTCTGTACTGGTTCGTGTAGATTTGGGCTGAGTTTCTCAAAACACACTACCAC 458
Qy      120  ValLysArgValTrpValSerAsnProSerTrpProAsnHis-LysSerValPheAsnse 139
Db      459  CAAAGTGTCAATTACATTCCTCAAAACCAACTGGGGGAACCATCCCAAGATTTTCAACTT 518
Qy      139  rAlaGlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPh 159
Db      519  GGCAGGCTTGTCTGTGGAGTATTTCCGCTACTATGATCTCTGCAACCCCGTGTGACTTT 578
Qy      159  eAspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisG1 179
Db      579  CAAAGCTTGTCTGAGGATCTTGGCGCTGACCATCTCGAGCTATTGTTCTTCTCATGC 638
Qy      179  yCysCysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaG1 199
Db      639  ATGTTACACATCCCACTGGAGTTGACCAACCTCTGAACAGTGGAAACAGATTCCGACA 698
Qy      199  nLeuSerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaAr 219
Db      699  GCTGATGAGATCTAAAAGCTTATTACCCCTTTTGTATAGTCATATCAGGGTTTGTCTAG 758
Qy      219  gGly---LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMethHisLysGluLe 238
Db      759  TGGTAGCCTTAAACACAGATGCACATCTGTCGTCATATTGTTGCTGATGGCGGTGAATG 818
Qy      238  uIleValAlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCy 258
Db      819  CTTGATAGTCAAAGTTAGCGCAAAATATCGGACTTTATGGGAGCGCTGTTGGTGCCT 878
Qy      258  sThrLeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAl 278
Db      879  TAGCATTGCTGCAAGTCAGCAGATGCGCTAGTAGTAAGGTTGAGAGCCAGGTGAAGCTTGT 938

```

```

Qy 278 aileAargAlaAenTyrSerAen-PropProAlaHisGlyAlaSerValValAlaThrIleL 298
Db 939 TGTGGGCCCATGTATTCGAGCCCCCTATTATGAGGACATCAATTGTGCCACCATTC 998

Qy 298 eUerAenAaspAlaLeuAargAlaIleTyrGluGlnGluLeuThrAaspMetAargGlnArgI 318
Db 999 TAAAGAGCAGTGATGTATACAACTCGGACCATCGAGCTGAAAGAAATGCTGACCGTA 1058

Qy 318 leGln--ArgMetAArgGlnLeuPheValAasnThrLeuGlnGluLysGlyAlaAasnArgAs 337
Db 1058 TAAAGCGGCATGCCCAACAGTATTGAGAGCTATACAGCTAGAGGTACACCTGGTTA 1118

Qy 337 pPheSerPheIleLysGlnAasnGlyMetPheSerPheSerGlyLeuThrLysGluG1 357
Db 1119 CTGGAGTCATATTATCAACAGATTGGGATGTTTACATTACTGGATTGAACAAGGAGCA 1178

Qy 357 nValLeuAargLeuAargGluGluPheGlyValTyrAlaValAlaSerGlyAargValAasnVa 377
Db 1179 AGTTGAATTCATGACCAAGAGTTCACATTACATGATGATGAGGAGAAATGAGCAT 1238

Qy 377 lAlaGlnMetTyrProAaspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1239 GCGAGTCTAAGTTGCAAGACAGTCCCTCACCTTGCCGATGCTATGCGATGCTGCAGTG 1296

RESULT 11
LOCUS AY815146 1502 bp mRNA linear HTC 23-FEB-2005
DEFINITION Schistosoma japonicum clone SJCHGC06163 unknown mRNA.
ACCESSION AY815146
VERSION AY815146.1 GI:56757418
KEYWORDS HTC.
SOURCE Schistosoma japonicum
ORGANISM Schistosoma japonicum
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeidida; Schistosomatoides; Schistosomatidae; Schistosoma.
REFERENCE 1 (bases 1 to 1502)
AUTHORS Han, Z.
TITLE The full-length cDNA sequences of Schistosoma japonicum genes
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1502)
AUTHORS Han, Z.
TITLE Direct Submission
JOURNAL Submitted (02-NOV-2004) Functional Genomic Department, Chinese
National Human Genome Center at Shanghai, 351 Guo Shoujing Road,
Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P.R.China
FEATURES
source
1..1502
/organism="Schistosoma japonicum"
/mol_type="mRNA"
/db_xref="taxon:6182"
/clone="SJCHGC06163"
53..1306
/note="similar to aspartate aminotransferase Complex with
Alpha-Methyl (45.6 kD) (XG861)"
/codon_start=1
/product="unknown"
/protein_id="AAW26878.1"
/db_xref="GI:56757418"
/translation="MPTKSPVTVTVIKRLSSHWNVKLPDPDAILGITAEYNRDTPQ
KINLGAYRDNDGKPFVPSVKEASLLAKNLNKEYAPISGIPQFCDLSIKULFTE
QSPRIKORCNATVQTSIGTALRIGGAFINEAQHIMWSPGWQHFIFVHSGLN
VHQRYDSTNCGPDIDCLSDLSKIPKHEVLLHACAHNPTGVDPSPDQWRKIGETM
KRDLIIPFDCAVQGFASGDDNNDAKAIKYFTDELNPFTLFTQSFKNMGLYGERVG
AFSLICSSDDEARCLSQIKILIRPMYSNPPIHGARTAAELMSNVDLRRKWLIDKSM
ADRIITKQSLKNGLIERGSDHSHITNQIMFCFSGLNFSQVEKUTNEYSIYLTIDK
GRISIALGSSKNVYLAAHAIHQVTK"

CDS
8,556-82 Length: 1502
807,50 Matches: 161
57.4% Conservative: 65

ALIGNMENT SCORES:
Pred. No.: 8,556-82 Length: 1502
Score: 807,50 Matches: 161
Percent Similarity: 57.4% Conservative: 65

```

```

Best Local Similarity: 40.9% Mismatches: 165
Query Match: 39.5% Indels: 3
DB: 4 Gaps: 2
US-10-673-786A-2 (1-396) x AY815146 (1-1502)

Qy 2 PheGluAenIleThrAlaAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPheArg 21
Db 107 TGGAAACCATGTTAAGCTGGGACCTCCAGATTCGAATTTTGGGAATTTACTGGAACCTACAAT 166

Qy 22 AlaAaspGluAargProGlyLysIleAasnLeuGlyIleGlyValTyrLysAaspGluThrGly 41
Db 167 CGTGACCAACCAATCCTCGAAAAATTAATCTAGGAGCTGGAGCCTATAGGATGATAATGGA 226

Qy 42 LysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAasnGluThr 61
Db 227 AAACCTTTTCTGTCCTGCCATCTGTCAAAGAGCGGAGTCACCTTTTATTAGCGAAAAATTTG 286

Qy 62 ThrLysAenTyrLeuGlyIleAaspGlyIleProGluPheGlyAargCysThrGlnGluLeu 81
Db 287 AATAAGGAGTATGCTCCAATCAGTGGCATTTCCCAATTTTTCGCGATTTTATCCATTAAAGCTT 346

Qy 82 LeuPheGlyLysGlySerAlaLeuIleAasnAaspLysAargAlaAargThrAlaGlnThrPro 101
Db 347 GCATTACTGAACAGTCTCCAGCAATCAAGATGCTGTGCAATGCTACAGTTCACACAAATA 406

Qy 102 GlyGlyThrGlyAlaLeuAargValAlaAaspPheLeuAlaLysAenThrSerValLys 121
Db 407 TCTGGCACTGCTCTCTAGCAATTTGGTGGAGCTTTTATTAAATGAATTCGCAGAACAAAAA 466

Qy 122 ArgValTrpValSerAenProSerTrpProAasnHisLysSerValPheAenSerAlaGly 141
Db 467 CACATTTGGATGCGCTGCACCTACTTGGGGCAATCATATACCCATTTTCGTGCATATGGA 526

Qy 142 LeuGluValAargGluTyrAlaTyrTyAspAlaGluAasnHisThrLeuAaspPheAaspAla 161
Db 527 TTGAATGTACATCAATATCGTTATTATGATAGTAATACATGTGGATTTGATTTATGTATGGA 586

Qy 162 LeuIleAasnSerLeuAasnGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCysCys 181
Db 587 TGTTGTCGGATTTGTCAAATAATTCCTAAAGGACATTTTGTTTTATTACATGCGATGTGA 646

Qy 182 HisAenProThrGlyIleAaspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeuSer 201
Db 647 CATAATCCTACAGAGTTGATTCATCATCTTTGATCAATGGAGGAAAAATTTGGAGAGATTATG 706

Qy 202 ValGluLysGlyTrpLeuProLeuPheAaspPheAlaTyrGlnGlyPheAlaAargGly--- 220
Db 707 AAATCAGTGATTTAATTCATTTTTCATTTGCTTATCAAGGATTCGCTTCAGGTGAC 766

Qy 221 LeuGluGluAaspAlaGluGlyLeuAargAlaPheAla-----AlaMetHisLysGluLeu 238
Db 767 ATTAATAATCATCGAAAGCAATAGCTTTATTTACAGACGAATTTGAATTTCCCTACTTTTA 826

Qy 239 IleValAlaSerSerTyrSerLysAenPheGlyLeuTyrAasnGluAargValGlyAlaCys 258
Db 827 TTCCTTACACAAAAGTTTCGCTAAAAATATGCGGATTTATATACGGGAAACGCTGTGGTGGT 886

Qy 259 ThrLeuValAlaAaspSerGluThrValAaspArgAlaPheSerGlnMetLysAlaAla 278
Db 887 AGTCTCTTTTCTTCATCAGATGAAGCGCAAGCTTTGTTTATCACAATTTAAATACTTT 946

Qy 279 IleAargAlaAasnTyrSerAenProAlaHisGlyAlaSerValValAlaThrIleLeu 298
Db 947 ATTCCACCAATGATAGTAAATCCACCAATTCACGGTGCACGCTATTGCTGCTGAATTAATG 1006

Qy 299 SerAenAaspAlaLeuAargAlaIleTyrGluGlnGluLeuThrAaspMetAargGluArgIle 318
Db 1007 TCTAATGTTGATTTACGTCGAAAAATGGCTCATTTGATTTGAAATCTATNTGGCTGATCGTATT 1066

Qy 319 GlnAargMetAargGlnLeuPheValAasnThrLeuGlnGluLysGlyAlaAasnAargPhe 338
Db 1067 ATAACAATGCGCCAAAGTTTTAAGANATGGTTTAAATCATCGAAGCAGGATCTCAACATGACTGG 1126

```

```

QY 339 SerPheIleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluInVal 358
Db 1127 TCACATATTACTAAATCAAAATGGAATGTTCTGCTTTTCTGGCTTAATCCATCACAAGTT 1186
QY 359 LeuArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAla 378
Db 1187 GAAATTAATTAACAAATGAAATATTCTATTATTATTTGACATAAGATGGAGCTATTAGTATGCT 1246
QY 379 GlyMetThrProAspAsnMetAlaProLeuCySGluAlaIle 392
Db 1247 GGATTTGCTCGGAAATGCGAAATATTAGCTCATGCTATT 1288

RESULT 12
LOCUS CR857622
DEFINITION Pongo pygmaeus mRNA; cDNA DKFZp459L1636 (from clone DKFZp459L1636).
ACCESSION CR857622
VERSION CR857622.1 GI:55726249
KEYWORDS HTC.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pongo.
1 (bases 1 to 2452)
Koehler,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Oeanger,A.,
Fobo,G., Han,M. and Wiemann,S.
The German cDNA Consortium
Direct Submission
Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA
sequencing consortium of the German Genome Project. This clone
(DKFZp459L1636) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp459L1636
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.
Location/Qualifiers
1..2452
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp459L1636"
/tissue_type="cortex"
/clone_lib="459 (synonym: pcor1). Vector pSport1_Sfi; host
DH10B; sites SfiIa + SfiIb"
/dev_stages="adult"
/note="aspartate transaminase precursor, mitochondrial
(Homo sapiens)"
1..2452
/gene="DKFZp459L1636"
103..1395
/gene="DKFZp459L1636"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH89897.1"
/db_xref="GI:55726250"
/translation="WALLHSGRALPGIAAAPHGLAAASARASRWTHVEMGPPDPI
LGVTEAFKRDNTNSKQNVGVAYRDNGKPYVLPVSRKAEQIAAKNLDKYLPIGGL
AEFCASAEALGENSEVYKSGRFVTQITSGTALRIGASFLQRFKFSRDLPLKP
SWGHHTPIFRDAGMOLQGVYVDPKTGFDFGTGVEDISKIPEQSVLLHACAHNPTG
VDPRPEONKEIATVVKNLPAFDMAYGFGASGDKDANAWRHFIHQGNVCLCQS
YAKNGLGKRGVATWCKDAEAKRVESQLKILIRPMYSPNPLNGARIAAILNTP
DLRKQLQLEVMKADRIIGMTQLVSNLKKGSTHWHQHTDQIGMFCFTGLKPEQVE
RLIKEFSIYMTKDGRI8VAGVTSNVGVLAHAHQVTK".

```

```

Alignment Scores:
Pred. No.: 2,95e-80 Length: 2452
Score: 797.00 Matches: 162
Percent Similarity: 58.3% Conservative: 66
Best Local Similarity: 41.4% Mismatches: 161
Query Match: 39.0% Indels: 2
DB: 4 Gaps: 2

US-10-673-786a-2 (1-396) x CR857622 (1-2452)

QY 4 AsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPheArgAlaasp 23
Db 205 CATGTGGAAATGGGACCTCCAGATCCCATCTGGGAGTCACTGAAGCCCTTTAAGGGAC 264
QY 24 GluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThrGlyLysThr 43
Db 265 ACCAATAGCAAAAGATGAACTGGAGTGTGTGCTACCGGATGATGATACCGAAAGCCT 324
QY 44 ProValLeuThrSerValLysAlaGluInTyrLeuLeuGluAsnGluThrThrLys 63
Db 325 TACGTGCTGCTAGCGTCCGCAAGCAGAGCCAGATTGCCGCAAAAATTTGGACAAG 384
QY 64 AsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGluGluLeuPhe 83
Db 385 GAATACCTGCCCATTTGGGGACTGGCTGAATTTTCAAGGCATCTGCAGAACTAGCCCTG 444
QY 84 GlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThrProGlyGly 103
Db 445 GGTGAGAACACGGAAGTCTTGAAGAGTGGCGGTTTGTCTACTGTGCAGACCACTTCTGA 504
QY 104 ThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerValLysArg--- 122
Db 505 ACTGAGCCTTAAGGATCGGAGCCAGTTTCTGCAAGATTTTAAAGTTTCAGCCGAGAT 564
QY 123 ValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerIleGlyLeu 142
Db 565 GTCTTTCTGCCCAACCATCTCGGGAAACACACACACCATCTTCAGGAGATCGGCATG 624
QY 143 GluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAspAlaLeu 162
Db 625 CAGCTACAGGTTATCGGTATTATGACCCCAAGACTTGGGTTTGTGACTTTCACAGGTGT 684
QY 163 IleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysHis 182
Db 685 GTGGAGGATATTTCAAAAATACCAGACAGAGTGTCTTCTTCTGCTGCTGCGCCAC 744
QY 183 AsnProThrGlyIleAspProThrLeuGluInTyrGlnThrLeuAlaGlnLeuSerVal 202
Db 745 AATCCACGGAGTGGACCCGCTCCGGAACAGTGAAGAAATAGCAACAGTGGTGAAG 804
QY 203 GluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly--Leu 221
Db 805 AAAAGGAATCTCTTTGCGTCTTTGACATGCCCTTACCAAGGCTTTGCCAGTGGTGATGT 864
QY 222 GluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleValAla 241
Db 865 GATAAGGATGCTGGGCTGTGGCCACTTCATCGAACAGGCGATTAAATGTTGTCTCTGC 924
QY 242 SerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeuVal 261
Db 925 CAATCATATGCCAAGAACATGGGCTTATGCTGAGCGTGTAGGAGCCCTTCACTATGCTC 984
QY 262 AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAla 281
Db 985 TGCAAGATGGGATGAAGCCAAAGGTAGAGTACAGTTCAGATTGAAAGTCTTGATCTGCTCC 1044
QY 282 AsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsnAsp 301
Db 1045 ATGTATTCCACCCTCCCTCAATGGGCGCGGATGTGCTGCTGCCATCTTGAACACCCCA 1104
QY 302 AlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArgMet 321
Db 1105 GATTTCGAAACAAATGGCTGCAAGAGTGAAGGAGCATGAAAGGAGCCGCGCATTCATTTGCATG 1164

```



```

Qy 322 ArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPheille 341
|||
Db 1165 CGGACTCAACTGGTCTCCAACTCAAGAGAGGGTTCCACCCACAACTGGCAACACATC 1224
|||
Qy 342 lIeYsGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeu 361
|||
Db 1225 ACCGACCAAAATTTGGCATGTTCTTTCACAGGGCTAAAGCCTGAACAGGTGGAGCGGCTG 1284
|||
Qy 362 ArgGluGluPheGlyValAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
|||
Db 1285 ATCAAGGAGTTCTCCATCTACATGACAAAGAGTGGCCCATCTCTGTGGCAGGGGTACCC 1344
|||
Qy 382 ProAspAsnMetAlaProLeuCysGluAlaIle 392
|||
Db 1345 TCCAGCAACGTGGGTACTCTTGCCCATGCCATT 1377
|||

RESULT 13
AW221110
LOCUS EST297579 tomato fruit mature green, TAMU Lycopersicon esculentum
DEFINITION cDNA clone CLEF12H11, mRNA sequence.
ACCESSION AW221110
VERSION AW221110.1 GI:5532794
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 468)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S.,
Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
Contact: CUGI
Clemson University
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
Location/Qualifiers
1..468
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLEF12H11"
/tissue_type="fruit pericarp"
/dev stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/clone_lib="tomato fruit mature green, TAMU"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLEF - Fruit were tagged at the 1cm stage and
harvested 3-5 days prior to ripening. Fruit were cut in
half to verify the seeds were indeed 'immature' and the
seeds and locules were discarded prior to freezing the
pericarp"
ORIGIN
Alignment Scores:
Pred. No.: 4,04e-81 Length: 468
Score: 795.00 Matches: 149
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 38.9% Indels: 0
DB: 1 Gaps: 0

US-10-673-786A-2 (1-396) x AW221110 (1-468)

```

* Qy 52 AlaGluGlnTyrLeuLeuGluAsnGluThrThrLysAsnTyrLeuGlyIleAspGlyIle 71

```

Db 2 GCTGAACAGATATCTGCTCGAAATGAAACCAACCAAAATTTACCTCGGCAITTCGCGGATC 61
|||
Qy 72 ProGluPheGlyArgCysThrGlnGluLeuPheGlyLysGlySerAlaLeuIleAsn 91
|||
Db 62 CTTGAATTTGGTGGCTGCATCTCAGGAATCTGCTTTGGTAAAGGTAGGCGCTCATCAAT 121
|||
Qy 92 AspLysArgAlaArgThrAlaGlnThrProGlyGlyThrGlyAlaLeuArgValAlaAla 111
|||
Db 122 GACAAACGTCTCGCAGCGCACAGACTCCGGGGGCGACTGCGCGCACTACGCGTGGCTGCC 181
|||
Qy 112 AspPheLeuAlaLysAsnThrSerValLysArgValTrpValSerAsnProSerTrpPro 131
|||
Db 182 GATTTCCTGCGCAAAATAACAGCGTTAAGCGTGTGGGGTGAGCAACCCCAAGCTGGCGC 241
|||
Qy 132 AsnHisLysSerValPheAsnSerAlaGlyLeuGluValArgGluTyrAlaTyrTyAsp 151
|||
Db 242 AACCATTAAGAGCGCTCTTAACCTCTGCAGGCTCGGAAGTTCGTGAATACGCTTATTATGAT 301
|||
Qy 152 AlaGluAsnHisThrLeuAspPheAspAlaLeuIleAsnSerLeuAsnGluAlaGlnAla 171
|||
Db 302 GCGGAAATCACACTCTTGACTTCGATGCATGATTAACAGCCTGAATGAAGCTCAGGCT 361
|||
Qy 172 GlyAspValValLeuPheHisGlyCysCysHisAsnProThrGlyIleAspProThrLeu 191
|||
Db 362 GCGACGTAAGTGTGTTCCATGCTGCTGCATAACCCACCGGGATCGACCCCTACGCTG 421
|||
Qy 192 GluGlnTrpGlnThrLeuAlaGlnLeu 200
|||
Db 422 GACATATGGCAACACTGGGCACACTA 448
|||

RESULT 14
CR860094
LOCUS 2447 bp mRNA linear HTC 12-NOV-2004
DEFINITION Pongo pygmaeus mRNA; cDNA DKFZp469B0125 (from clone DKFZp469B0125).
ACCESSION CR860094
VERSION CR860094.1 GI:55731047
KEYWORDS HTC.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pongo.
1 (bases 1 to 2447)
Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
The German cDNA Consortium
Direct Submission
Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Agowa (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp469B0125) is available at the RZPD Deutsches
Reisourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469B0125
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.
FEATURES
Location/Qualifiers
1..2447
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp469B0125"
/tissue_type="kidney"
/clone_lib="469 (synonym: pkid1). Vector pSport1_Sfi; host
DH10B; sites SfilA + SfilB"
/dev stage="adult"
/note="aspartate transaminase precursor, mitochondrial
(Homo sapiens)"
1..2447

```

gene


```

/ gene = "DKFZp469B0125"
Join(103...1116,1116...1394)
/ gene = "DKFZp469B0125"
/ codon_start = 1
/ product = "hypothetical protein"
/ protein_id = "CAH92240.1"
/ db_xref = "GI:55731048"
/ translation = "WALLHSGRALPGIAAAPHGLAASARASSWTHVEMGPPDPPI
LCVTAERFDNTSKMKNLGVGAYRDDNGKPYVLPVSRVKAQIAAKNLDKEYLPTGGI
AEFCASAEALAGENSEVLKSGRFVTVOTISGTGALRIGASFLORFFKPSRDVLPKLP
SKWNTHTPFRDAGMQLQGYRYDPTKCGDFDFTGAVEDISKIPQSVLLHACAHNPTG
VAPPEQKATVYVKKNLFAFFDMAYGFSAGGDDKDAWVRHFIISGINVLCQS
YAKNGLGGERVGAFTWCKDAEKRVESQKILIRWYNNPPLNGARIAIAALNTP
DLRTKTLQEVGMADRIICMRTQLVSNLKGSTHWOHTIDQIGMFCFTGLKPRQVE
RLIKEFSYMTKDRISVAGVTSNVGVLAHAHQVTK"

```

ORIGIN

Alignment Scores:

Pred. No.:	6.03e-78	Length:	2447
Score:	777.00	Matches:	162
Percent Similarity:	58.3%	Conservative:	66
Best Local Similarity:	41.4%	Mismatches:	161
Query Match:	38.0%	Indels:	3
DB:	4	Gaps:	2

US-10-673-786A-2 (1-396) x CR860094 (1-2447)

```

QY 4 AsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPheArgAlaAsp 23
DB 205 CATGTGGAAATGGGACCTCCACATCCCATTTGGGAGTCACCTGAAGCCCTTTAAGAGGGAC 264
QY 24 GluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThrGlyLysThr 43
DB 265 ACCAATAGCAAAAGATGATCTGGAGTTGGTGCTACCGGATGATACCGAAAGCCT 324
QY 44 ProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGluThrThrLys 63
DB 325 TAGCTGCTGCTAGCGTCCGAAGCGAGAGCCAGATGGCCGCAAAATTTGGACAAG 384
QY 64 AsnTyrLeuGlyLysIleAspGlyLeuProGluPheGlyArgCysThrGlnLeuLeuPhe 83
DB 385 GAATACCTGCTCCATTTGGGGAGCTGGCTGAATTTTGCAGAGCATCTGCAGAGTAGCCCTG 444
QY 84 GlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThrProGlyGly 103
DB 445 GTTGAGAACGCAAGCTCTGAAGAGTGCCCGGTTGTCTGTCGTGCAGACCAATTTCTGA 504
QY 104 ThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerValLysArg--- 122
DB 505 ACTGGAGCCTTAAGGATCGGACCGCATTTTCTGCAAGATTTTAAAGTTCAGCCGAGAT 564
QY 123 ValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerIleGlyLeu 142
DB 565 GTCTTTCTCCCAAAACCATCTCTGGGGAACACACACACCATCTTCAGGGATGCTGGCATG 624
QY 143 GluValArgGluTyrAlaTyrTrpAlaGluAsnHisThrLeuAspPheAspAlaLeu 162
DB 625 CAGCTACAGAGGTATCGGTATTATGACCCCAAGACTTGGCGTTTTCGACTTCACGGTGT 684
QY 163 IleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCysHis 182
DB 685 GTGGAGGATATTTCAAAATACACAGCAGAGAGTGTCTTCTGTCATGCTCGCCGCCAC 744
QY 183 AsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeuSerVal 202
DB 745 AATCCACGGGAGTGACCCCGCTCCGGAACAGTGGGAAGAAATAGCAACAGTGGTGAAG 804
QY 203 GluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly---Leu 221
DB 805 AAAAGGAATCTCTTTGCGTCTTCTGACATGGCGCTACCAAGGCTTTGCCAGTGGTGAAGT 864
QY 222 GluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGlyLeuIleValAla 241

```

```

DB 865 GATAAGGATGCTGGGCTGTGGCCACATTCATCGAAACAGGGCATTAATGTTTGTCTCTCC 924
QY 242 SerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeuVal 261
DB 925 CAATCATATGCCNAGAACATGGCTTATATGCTAGCGGTAGGAGCCCTTCACTATGCTC 984
QY 262 AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAla 281
DB 985 TGCAAGATGCGGATGAAGCCAAAGAGGTAGAGTCACAGTTCAGAGATCTTGTATCGCTCC 1044
QY 282 AsnTyrSerAsnProAlaHisGlyAlaSerValAlaThrIleLeuSerAsnAsp 301
DB 1045 ATGTATTCCACCTCCCTCAATGGGCGCGGATGCTGCTGCCATCTCGAACACCCCA 1104
QY 302 AlaLeuArgAlaIleTyrGlnGluLeuThrAspMetArgGlnArgIleGlnArgMet 321
DB 1105 GATTTCGAAA-ACATGGCTGCNAGAGTGAAGGCGATGGCCGCGCATCATTTGGCATG 1163
QY 322 ArgGlnLeuPheValAsnThrLeuGlnGlyLysGlyAlaAsnArgAspPheSerPheIle 341
DB 1164 CGGACTCAACTGGTCTCCAACTCAAGAAGAGGGGTTCACCCACCAACTGGCAACACATC 1223
QY 342 IleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGlnValLeuArgLeu 361
DB 1224 ACCGACCAAAATGGCATGTTCTGTTTCACAGGGCTAAAGCCTGAACAGGTGGAGGGCTG 1283
QY 362 ArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
DB 1284 ATCAAGGAGTTCTCCATCTACATGACAAGGATGGCCGATCTCTGTGCGAGGGTCAAC 1343
QY 382 ProAspAsnMetAlaProLeuCysGluAlaIle 392
DB 1344 TCCGCAACGTGGGTACTCTTGCCTCATTCGCATT 1376
RESULT 15
CR616132 1496 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0CAP003YO15 of Thymus of Homo sapiens
(human).
DEFINITION CR616132
ACCESSION CR616132 GI:50496939
VERSION CR616132.1
KEYWORDS HTC; CNSLT cDNA.
SOURCE Full-length cDNA libraries and normalization
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarctostylis; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1496)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
REFERENCE 2 (bases 1 to 1496)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1..1496
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP003YO15"
/tissue_type="Thymus"
/plasmid="pCMVSPORT_6"

```


GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 03:10:41 ; Search time 390 Seconds
(without alignments)
1804.909 Million cell updates/sec

Title: US-10-673-786A-2
Perfect score: 2045
Sequence: 1 MFENTRAAPADPILGLDLF.....VAGMTPDNAPICEAIVAVL 396

Scoring table: BLOSUM62

Qgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlp
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSSWER_spool/US10673786/runat_14032006_135613_22673/app_query.fasta.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02p -USER=US10673786 @CGN 1.1 512 @runat_14032006_135613_22673
-NCPU=6 -ICPU=3 -NO MAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRRADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PTCUT COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2045	100.0	1331	3	Sequence 23, Appl
2	1931	94.4	1209	3	Sequence 4902, Ap
3	1325	64.8	1830121	3	Sequence 1, Appli
4	1325	64.8	1830121	3	Sequence 1, Appli
5	1325	64.8	1830121	3	Sequence 1, Appli
6	982	48.0	1242	3	Sequence 12428, A
7	982	48.0	1329	3	Sequence 252-991A-12428
8	982	48.0	24417	2	Sequence 12162, A
9	955	46.7	579	3	Sequence 4954, Ap

c	10	909.5	44.5	1509	3	US-09-252-991A-12968	Sequence 12968, A
	11	909.5	44.5	1632	3	US-09-252-991A-12507	Sequence 12507, A
	12	908	44.4	702	3	US-09-543-681A-3805	Sequence 3805, Ap
	13	872.5	42.7	100848	3	US-09-596-002-39	Sequence 39, Appl
	14	871.5	42.6	1209	3	US-09-540-236-315	Sequence 315, App
	15	844.5	41.3	1338	3	US-09-489-039A-1301	Sequence 1301, Ap
	16	843.5	41.2	1260	3	US-09-543-681A-3123	Sequence 3123, Ap
	17	830.5	40.6	1284	3	US-09-328-352-769	Sequence 769, App
c	18	824.5	40.3	1185	3	US-09-489-039A-1274	Sequence 1274, Ap
	19	794	38.8	2440	3	US-09-976-594-429	Sequence 429, App
	20	759	37.1	1941	3	US-09-220-132-65	Sequence 65, Appl
	21	759	37.1	1941	3	US-09-814-015A-35	Sequence 35, Appl
	22	759	37.1	1942	3	US-09-949-016-1216	Sequence 1216, Ap
	23	759	37.1	2056	3	US-09-566-921-56	Sequence 56, Appl
c	24	726.5	35.5	1230025	3	US-09-198-452A-1	Sequence 1, Appli
c	25	726.5	35.5	1230230	3	US-09-438-185A-1	Sequence 1, Appli
	26	663	32.4	1127	3	US-09-533-559-59	Sequence 59, Appl
	27	546	26.7	912	3	US-09-248-796A-3374	Sequence 3374, Ap
	28	514	25.1	988	3	US-09-533-559-145	Sequence 145, App
	29	510	24.9	426	3	US-09-543-681A-3717	Sequence 3717, Ap
	30	494.5	24.2	900	3	US-09-533-559-4412	Sequence 4412, Ap
	31	486.5	23.8	867	3	US-09-248-796A-3375	Sequence 3375, Ap
	32	365	17.8	423	3	US-09-252-991A-12380	Sequence 12380, A
	33	359.5	17.6	1517	3	US-09-801-874-1	Sequence 1, Appli
c	34	343	16.8	668	3	US-09-533-559-4436	Sequence 4436, Ap
	35	329	16.1	207	3	US-09-489-039A-4873	Sequence 4873, Ap
	36	325.5	15.9	646	3	US-09-328-111-658	Sequence 658, App
	37	322	15.7	423	3	US-09-252-991A-12333	Sequence 12333, A
	38	321.5	15.7	726	3	US-09-248-796A-3377	Sequence 3377, Ap
	39	317.5	15.5	658	3	US-09-533-559-4632	Sequence 4632, Ap
	40	313	15.3	524032	3	US-09-949-016-16928	Sequence 16928, A
	41	313	15.3	524032	3	US-09-949-016-16929	Sequence 16929, A
	42	313	15.3	524032	3	US-09-949-016-16930	Sequence 16930, A
	43	313	15.3	524032	3	US-09-949-016-16931	Sequence 16931, A
	44	313	15.3	529885	3	US-09-949-016-14340	Sequence 14340, A
	45	313	15.3	529885	3	US-09-949-016-14341	Sequence 14341, A

ALIGNMENTS

RESULT 1

US-08-985-908-23
; Sequence 23, Application US/08985908
; Patent No. 6004773

GENERAL INFORMATION:

APPLICANT: MASAYUKI ARAKI, MASAKAZU SUGIMOTO, YASUHIKO YOSHIHARA, AND TSUYOSHI NA
TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,908
FILING DATE: 05-DEC-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 8-325659
FILING DATE: 05-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 23:

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1331 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; STRAIN: JM109
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10...1197
; US-08-985-908-23

Alignment Scores:
Pred. No.: 5,31e-240 Length: 1331
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-673-786A-2 (1-396) x US-08-985-908-23 (1-1331)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 10 ATGTTTGAGAACATTACCGCGCTCTGCCGACCGGATTCTGGGCTGGCGGATCTGTTT 69
Qy 21 ArgAlaAspGluArgProGlyValIleAsnLeuGlyIleGlyValTyrlsAspGluThr 40
Db 70 CGTCCGATGAACGCTCCCGGCAAAATTAACTCGGGATTGGTGTCTATAAAGATGAGACG 129
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 130 GGCAGAACCCCGGTACTACCGAGCTGAAAGGCTGAACAGTATCTGCTCGAANAATGAA 189
Qy 61 ThrThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 190 ACCACCAAAATTAACCTCGGATTCAGCGCATCCCTGTAATTTGGTCGTCACACAGAA 249
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaAaThrAlaGlnThr 100
Db 250 CTGCTGTTTGGTAAGGTAGCGCCCTGATCAATGACAAACGTGTCGACGCGCACAGACT 309
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 310 CCGGGGGGCACTGGGCGACTAGCGTGGCTGCCGATTTCTTGGCAAAAATACAGCGTT 369
Qy 121 LysArgValTrpValSerAsnProSerTrpProAenHisLysSerValPheAsnSerAla 140
Db 370 AAGCGTGTGTGGTGAGCAACCCCAAGCTGGCGCAACCAATAAGAGCGTCTTTAACTCTGCA 429
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAenHisThrLeuAspPheAsp 160
Db 430 GGTCTGGAAGTTCGTGAATACGCTATTATGATCGGAAATACACTCTTTGACTTCGAT 489
Qy 161 AlaLeuIleAenSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 490 GCACGTATTACAGCCTGAAAGAGCTCAGGCTGGCGAGTAGTGTCTTCATGGCTGC 549
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 550 TGCCATAACCCCAACCGGTATCGACCTACGCTGGCAACCAATGCAACACATCGCACACTC 609
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 610 TCCGTTGAGAAGGCTGTTTACCGCTGTTGACTTCGCTTACCAAGGGTTTTCGCCGTGT 669
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
Db 670 CTGGAAGAAGATGCTGAAGAGGCTGCGCGCTTTCGCGGCTATGCAATAAAGAGCTGATGTT 729
Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1331 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Klebsiella pneumoniae
; STRAIN: ATCC 35962
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10...1197
; US-09-489-039A-4902

Alignment Scores:
Pred. No.: 4,07e-226 Length: 1209
Score: 1931.00 Matches: 370
Percent Similarity: 96.5% Conservatives: 12
Best Local Similarity: 93.4% Mismatches: 14
Query Match: 94.4% Indels: 0
DB: 3 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-489-039A-4902 (1-1209)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 19 ATGTTTGAGAACATTACCGCGCGCCCGCCAGCCGATTTAGGTCTGGCGGATCTGTTT 78
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrlsAspGluThr 40
Db 79 CGTCCGATGACCGCCCTGAAAAAATAAATTAACCTCGGAATTTGGTGTTCACAGATGAAC 138
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 139 GGTAAAAAGCCCGCTTCTGACGAGCGTCAAAAAGCAGACGATATCTGCTGAAAAATGAA 198
```

61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 199 ACCGCTAAAACATATCTGGCATCGATGTTATCTCTGAATTTGGTCTGCACCCAGGAG 258
 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 259 CTGCTGTTTCGGTAAGCAATGCGATTATCGCTGATAAAGCGCCCGCCGCGCAGACA 318
 101 ProGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 319 CCGCGCGTACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 378
 121 LysArgValTrpValSerLysProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 379 AAACGTGTGTGGTAAGTAATCCGAGTGGCGGACCAATAAAGCGTATTCACCTCTGCC 438
 141 GlyLeuGluValArgGlyTrpAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 439 GGGCTGGAAGTGGCGAATACGCATACGACGCGCGCTAACCCAGCGCTGGACTTTGAT 498
 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 499 GGTCTGCTGCCAGCTGAACGAGCCAGCGCGGCGAGCTGGTACTGTTCACCGCTGC 558
 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 559 TGCCACAACCCGACCGTATCGATCGACGCTCGATCGATCGATCGATCGATCGATCG 618
 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 619 TCGGTGGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 678
 221 LeuGluGluAspAlaGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
 679 CTGGAAGACAGTCTGAAGGCTGACCGCTTTTGTCTGTCTACATAAAGAGCTGTCTGC 738
 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 739 GCCAGCTCTACTCGAAGAACTTTGGCTGTGACACAGCGCGCTGGCGCTGCACTCTG 798
 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 799 GTCCGCGCGGATCAGGAGACTGTAGACCGCGCTTCAGTCAGATGAAGTCGGTGTCCGC 858
 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 859 GCCAATCTCGAACCCTGCGCATGCGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 918
 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 919 GATCGCTACCGGCAATCTGGGAGCAGGAACCTGACCGATATGCGCCAGCGATCCAGCGT 978
 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
 979 ATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1038
 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 1039 ATCAGCCAGCAAGCGCATGCTTCATTCACGCGCGCTGACTAAAGAGCAGGTGTGCGC 1098
 361 LeuArgGluGluPheGlyValTrpAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 1099 CTGCGTGAAGAGTTCGCCATCTATGGGTAGCTTCGCGACGATCAACGTGGCGCGGATG 1158
 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 1159 ACGCTGCAATATGCGCGCGCTGTGCGAAGCCATCGTCCGCTACTG 1206

RESULT 3
 US-09-557-884-1

; Sequence 1, Application US/09557884
 ; Patent No. 6506581
 ; GENERAL INFORMATION:

APPLICANT: Fleischmann et al.
 TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof
 NUMBER OF SEQUENCES: 1
 CORRESPONDENCE ADDRESS:
 ADDRESS: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: MD
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3 1/2 inch diskette
 COMPUTER: Dell Pentium
 OPERATING SYSTEM: MS DOS v6.22
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/557,884
 FILING DATE: 25-Apr-2000
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/476,102
 FILING DATE: JUN-5-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Michelle S. Marks
 REGISTRATION NUMBER: 41,971
 REFERENCE/DOCKET NUMBER: PB186P3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 301-309-8504
 TELEFAX: 301-309-8439
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1830121 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-557-884-1
 Alignment Scores:
 Pred. No.: 7,18e-147 Length: 1830121
 Score: 1325.00 Matches: 248
 Percent Similarity: 77.8% Conservative: 60
 Best Local Similarity: 62.6% Mismatches: 88
 Query Match: 64.8% Indels: 0
 DB: 3 Gaps: 0
 US-10-673-786A-2 (1-396) x US-09-557-884-1 (1-1830121)
 QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 DB 1684147 ATGTTTGAACATATCAAGCGCGCACCGCATCCATCTTAGGCTTAGCGAAGCATTT 1684206
 QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 DB 1684207 AAATCCGAACCTCGGAAAATAAATCAATTGGGTATTGGGTTTATAAGATGCGCAA 1684266
 QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
 DB 1684267 GGCACACCCCAATTTATGACGCGGTAAAGAACGCGAAGAACGATTTATTGATAAGGAA 1684326
 QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 DB 1684327 AAACCAAGAATATCTGACTATCGATGCTATTCGGATTTATACGAACAAACAAAGCA 1684386
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 DB 1684387 CTCCTTTTCGGTAAGATTTCTGAAGTCATCAATCTAATCGGCAAGAACAGTACAAAGT 1684446
 QY 101 ProGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 DB 1684447 TTAGCGGAAACAGGTGCTATTACGATTCGGCAGAAATTTATTAAACGCCAACTAAAGCA 1684506

```

Qy 121 LysArgValTrpValSerAsnProSerTrpProHemHisLysSerValPheAsnSerAla 140
Db 1684507 CAAATGTTGGTTCAGCACCACTTGGCCAAACCAATGCGATTTCAATGCTGTC 1684566
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTrpAlaGluAsnHisThrLeuAspPheAsp 160
Db 1684567 GGTATGACCAATTCGGAATATCGTTATATGATGCTGAACCAAGCCCTTGATGGGAA 1684626
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 1684627 CACTATTAGAGATTAAAGCCACGACGAGCGATGTTGCTTTTACACGGTGT 1684686
Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 1684687 TGCCATAATCCGACTGGTATTGACCTACTCCAGAACCAATGCAAGATTAGCCGCACTT 1684746
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 1684747 TCAGCTAAATATGGTGGTGGCTCTTTGACTTTGCTTATCAAGGTTTACCCACGGA 1684806
Qy 221 LeuGluGluAspAlaGluGluLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 1684807 TTAGATGAAGATGCTTATGGTTAGCTGCTTTTCAGCAGCAACCAACCAAGATTATTAGT 1684866
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 1684867 GCGAGTTCATTCGAAATATGCTTATGTTTATATGAATGAACTGTTGGTGCATTTACCCCTT 1684926
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 1684927 GTGGCAGAAATGCGAAGATTGCATCAACCTCATTAACCCCAAGTAAATCAATTTATTCGC 1684986
Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 1684987 ACCCTATCTCTTACCCAGCTTCTCAGCGGGGCGACCGTAGCAGACAGTATTAAATGAT 1685046
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 1685047 GCTCAACTTCGCCAAGATGGGAAATCAATTAAGTAAATGCGTGAACGATCAAAAAA 1685106
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 1685107 ATGCGTCACTATTTCGTTTACGTTTATTAAGAAATATGCTGCGAGAACCAAGATTTCAGCTTT 1685166
Qy 341 IleIleLysGlnAsnGlyMetPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1685167 ATCATTAACCAACCGATGATGTTTCTTTCAGTGGATTACAGGGGAACAAGTGGATCGT 1685226
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAlaValAlaGlyMet 380
Db 1685227 TTAATAAATGAATTTGCCATTTAGCTGTTCTGTTCTGCTGCTATCAACGATAGCTGGAATC 1685286
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1685287 ACAGAAGATAATATTCGTTATCTATGTGAAGATATCGTGAAGTACTT 1685334

```

RESULT 4

US-09-643-990A-1
Sequence 1, Application US/09643990A
Patent No. 6528289

GENERAL INFORMATION:

APPLICANT: Robert D. Fleischmann

Mark D. Adams

Owen White

Hamilton O. Smith

J. Craig Venter

TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

```

; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB186P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

```

Alignment Scores:

```

Pred. No.: 7,18e-147 Length: 1830121
Score: 1325.00 Matches: 248
Percent Similarity: 77.8% Conservative: 60
Best Local Similarity: 62.6% Mismatches: 88
Query Match: 64.8% Indels: 0
DB: 3 Gaps: 0

```

US-10-673-786A-2 (1-396) x US-09-643-990A-1 (1-1830121)

```

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1684147 ATGTTGAACATATCAAGCGGCGCCAGCCGATCCAATCTTAGGCTTAGCGGAAGCATTT 1684206
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 1684207 AAATCCGAACTCGCAAAAATAAAATCAATTTGGGTATTGGCGTTTATAAAGATGCGCAA 1684266
Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 1684267 GGCACAAACCCCAATTTATGACCGCGGTAAAGAGCCGAAACGATTTATTGATAAGGAA 1684326
Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGlyArgCysThrGlnGlu 80
Db 1684327 AAAACCAAGAAATTTATCTGACTATCGATGGTATTGCGGATTTATACGACAAACAAGCA 1684386
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 1684387 CTCCTTTTCGTTAAAGATTCTGAAGTCTCATCAATCTAATCGAGCAAGACAGTACAAAGT 1684446
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 1684447 TTAGCGGGAACAGGTGCATTTACGATTCGCGGAGAAATTTATTAAACCCCAACCTAAAGCA 1684506
Qy 121 LysArgValTrpValSerAsnProSerTrpProHemHisLysSerValPheAsnSerAla 140
Db 1684507 CAAATGTTGGTTCAGCACCACTTGGCCAAACCAATGCGATTTCAATGCTGTC 1684566

```

Qy	141	GlyLeuGluValValArgGluTyrAlaTyTyrAspAlaGluAsnHisThrLeuAspPheAsp	160
Db	1684567	GGTATGACCAATTCTGTGAATATCGTTATTATGATGCTGAACCAAGCCCTTGATTGGGA	1684626
Qy	161	AlaLeuLeuLeuSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys	180
Db	1684627	CACCTATTAGAAGATTTAAAGCAAGCANGCAAGGCGATGTGTGCTTTTACACGGTGT	1684686
Qy	181	CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu	200
Db	1684687	TGCCAATAATCCGATGGTATTGACCTTACTCCAGAACATGGCAGAAATTAGCCGACATT	1684746
Qy	201	SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly	220
Db	1684747	TCAGCTAAAAATGGTTGGTTCACACTCTTTGACTTTTGCATTATCAAGGTTTAGCCACGGA	1684806
Qy	221	LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMethIleHisGlyLeuLeuIleVal	240
Db	1684807	TTAGATGAAGATGCTTATGGTTTACGGTCTTTTGCAGCAAAACCACCAAGAAATATTATG	1684866
Qy	241	AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu	260
Db	1684867	GCAGTTTCATCTCGAAAAAATCTTGGTTTATATATGAACAGTGTGGTGCAATTACCCCT	1684926
Qy	261	ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMethIleValAlaIleArg	280
Db	1684927	GTGCGAAAAATGCAGAAATTTGCATCAACCTCATTAACCCCAAGTAAATCAATTATTCG	1684986
Qy	281	AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn	300
Db	1684987	ACCCATTACTCTTAACCCAGCTTCTCAGCGGGGGCGGACCGTAGCAACAGTATTAAATGAT	1685046
Qy	301	AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg	320
Db	1685047	GCTCACTTCGCCNAGNATGGGAAATGANTTAACCTGAATGGGTGACGCATCAAAAA	1685106
Qy	321	MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe	340
Db	1685107	ATCGCTCACCTATTCTCGTTTCAGTTTATTAAGAATATGGTGCAGAACCAAGATTTCAGCTTT	1685166
Qy	341	IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg	360
Db	1685167	ATCATTTGAACAAACCGGTATGTCTTCTTCAGTGGATTTAACAGGGGAACAAGTGGATCGT	1685226
Qy	361	LeuArgGluGluPheGlyValTyrAlaValAlaLaserGlyArgValAsnValAlaGlyMet	380
Db	1685227	TTAAAAATGAATTTGCCATTATCGCTTTCGTTCTGCTGATCATCAACGTAGCTGAATC	1685286
Qy	381	ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu	396
Db	1685287	ACAGAAGATTAATTCGCTACTCTATGTGAAGTATCGTGAAGACTCTT	1685334

RESULT 5

```

US-10-158-865-1
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, "Frage"
; Patent No. 6846651
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2G1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1

```



```

LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)

```

:	OTHER INFORMATION:	n equals a,t,c, or g
:	FEATURE:	
:	NAME/KEY:	misc_feature
:	LOCATION:	(140398)..(140398)
:	OTHER INFORMATION:	n equals a,t,c, or g
:	FEATURE:	
:	NAME/KEY:	misc_feature
:	LOCATION:	(142750)..(142750)
:	OTHER INFORMATION:	n equals a,t,c, or g
:	FEATURE:	
:	NAME/KEY:	misc_feature
:	LOCATION:	(145058)..(145058)
:	OTHER INFORMATION:	n equals a,t,c, or g
:	FEATURE:	
:	NAME/KEY:	misc_feature
:	LOCATION:	(145171)..(145171)
:	OTHER INFORMATION:	n equals a,t,c, or g
:	FEATURE:	
:	NAME/KEY:	misc_feature
:	LOCATION:	(145942)..(145942)
:	OTHER INFORMATION:	n equals a,t,c, or g
:	FEATURE:	
:	NAME/KEY:	misc_feature
:	LOCATION:	(147197)..(147197)
:	OTHER INFORMATION:	n equals a,t,c, or g
:	FEATURE:	
:	NAME/KEY:	misc_feature
:	LOCATION:	(150841)..(150841)
:	OTHER INFORMATION:	n equals a,t,c, or g
:	FEATURE:	
:	NAME/KEY:	misc_feature
:	LOCATION:	(152500)..(152500)
:	OTHER INFORMATION:	n equals a,t,c, or g
:	FEATURE:	
:	NAME/KEY:	misc_feature

Alignment Scores:					
Pred. No.:	7.18e-147	Length:	1830121		
Score:	1325.00	Matches:	248		
Percent Similarity:	77.8%	Conservative:	60		
Best Local Similarity:	62.6%	Mismatches:	88		
Query Match:	64.8%	Indels:	0		
DB:	3	Gaps:	0		

US-10-673-786A-2 (1-396) x US-10-158-865-1 (1-1830121)					
Qy	1	MetPhoGluAsnIleThrAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe	20		
Dd	1684147	ATGTGGTGAACATATCAAGCGGCACCAGCCGATCCAATTCTTAGGCTTTAGCGAAGCATTT	1684206		
Qy	21	ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrlYsAspGluThr	40		
Dd	1684207	AATTCGAAACTCGCAAAAATAATCAATTTGGGTATTGGCGTTTATAAGATCGCAA	1684266		
Qy	41	GlyLyThrProValLeuThrSerValLysLysAlaGluGlnTyrlLeuGluAsnGlu	60		
Dd	1684267	GGCACAACCCCAATTATGCACGCGGTAAAAAGAAGCGAAAAAACGATTATTTCGATAAGGAA	1684326		
Qy	61	ThrThrylAsnTyrlLeuGlyIleAspGlyIleProGluPheGlyArgCystrGlnGlu	80		
Dd	1684327	AAAACCAAGAATTTATCTGACTATCGATGTGTATTGGCGATTATAACGAACAACAAAGCA	1684386		
Qy	81	LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaargThrAlaGlnThr	100		
Dd	1684387	CTCTTTTTCGGTAAAGATTCTGNAGTCATCCAATCTAATCGAGCAAGAACACGACACAAAGT	1684446		
Qy	101	ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal	120		
Dd	1684447	TTAGCGGGAACAGGTGCATTAGCATTTGCCGAGCAATTTATTAAACGCCCAACTAAAGCA	1684506		
Qy	121	LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla	140		
Dd	1684507	CAAAATGTTTGWGTACGACACCAACTTGGCCAAACCAACAATGCCGATTTTCAAGCTGTC	1684566		

141 GlyLeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 1684567 GGTATGACCATTCGTGAATATCGTTATTATGATGCTGACGCAAGCAAGCCCTTGTATGGGAA 1684626
 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 1684627 CACTATTAGAAGATTAAAGCAAGCAAGCAAGCGATGTGTGCTTTTACACGGTTGT 1684686
 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 1684687 TGCCATAATCCGACCTGGTATTGACCTCTCCAGAACATGGCAAGATTAGCCGACATT 1684746
 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 1684747 TCAGCTAAATAATGGTTGGTGGCACTCTTTGACTTTTCAAGTTTAGCCCAACGA 1684806
 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuVal 240
 1684807 TTAGATGAAGATGCTTATGTTTACGTTTTCAGCAAAACCAACCAAGAAATTATTAGTG 1684866
 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 1684867 GCGAGTTCAATCTCGAAATAATTTGGTTTATATAATGAACGGTGTGGTGCAATTTACCCCT 1684926
 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 1684927 GTGGCAGAAAATGCAAGAAATTCATCAACCTCTATTAACCAAGTAATCAATATTATCGC 1684986
 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 1684987 ACCCTATACCTTAACCCAGCTTCTACGGCGGGCGACCGTAGCAACAGTATTAAATGAT 1685046
 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 1685047 GCTCAACTTCGCAAGAAATGGGAAAATGAATTAACCTGAATCGTGAACGCATCAAAAAA 1685106
 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
 1685107 ATGGTCACCTATTCTTCAGTTATTAAAGAAATATGTTGTCAGAACAAAGATTTCAGCTTT 1685166
 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 1685167 ATCATTTGAACAAAACGGTATGTTTCTTCAGTGGATTAAACAGGGAACAAGTGGATCGT 1685226
 361 LeuArgGluLysPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 1685227 TTAATAAATGAATTTGCCATTTACGCTGTTCTGTTCTGTTCTGTTATCAACAGTAGCTGAATC 1685286
 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 1685287 ACAGAAGATATATTCGCTATCTATGTGAAGATATCGTGAAGATCTT 1685334

RESULT 6
 US-09-252-991A-12428
 ; Sequence 12428, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196, 136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 12428
 ; LENGTH: 1242
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-12428
 Alignment Scores:
 Pred. No.: 5, 81e-110 Length: 1242
 Score: 982.00 Matches: 190
 Percent Similarity: 65.7% Conservative: 70
 Best Local Similarity: 48.0% Mismatches: 136
 Query Match: 48.0% Indels: 0
 DB: 3 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-252-991A-12428 (1-1242)
 QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 52 CTGTTTCTGCGTGAATGCAATGCAACCGGTGATCCATCTTGGCTGAACGAAGCCTTC 111
 QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 Db 112 AACGCCGATACCCGTCGCGGCAAGATCAACCTGGCGGTGGCGGTGACTACACGAGGAG 171
 QY 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
 Db 172 GGGCGCATTCGCTGTCGTCGCGGTGCGAGGCTGCGAGAGAGCCGCGCATCGAGGCCAC 231
 QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 232 GCGCGCGCGGTACTCGCATCGAAGGATCGCGCGCTACGACCAGGCGGTACAGAAA 291
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuLeuAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 292 CTGCTGTTTCGGTAACGAGTCCGAGCTGCTGGCGCGCGCGGTGCTGCTACGACCCAGGCC 351
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 352 GTCGCGCGCACCGCGCGCTCAAGCTCGCGCGCGACTTCTCAAGCGCTGCTGCCCGAC 411
 QY 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 Db 412 GCCACCGTGGCCATCAGCGACCCGAGCTGGGAAAACACCGCGCACTGTTTCAAGCGGCC 471
 QY 141 GlyLeuGluValArgGlyTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 Db 472 GGTTCCTCCGCTGACAACTACCGCTATTACGACCGCGCGCGCAATGGCGTGAACCGCGCC 531
 QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 532 GGCCTGCTGGAGACGTAATGCCCTGCGCGCGCGCTTCGATCGTGTCTGTCGACGCTGC 591
 QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 Db 592 TGCCATATCCGACCGCGCTCGATCTCGAGCTGGACGACTGGAAACAGGTGCTGGACGCTG 651
 QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 652 CTCAGGCAAGGCGCGCTGCTGCTGCAATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 711
 QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 Db 712 ATCGAGGAAGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 771
 QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValAlaCysThrLeu 260
 Db 772 TCCAGCTGTTCTCCAAATCTTCTGCTCTATGCGAAGCGGTGCGCGCCCTCTCGATC 831
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 832 GTAACGAAAGCGCGCATGAATCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891
 QY 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 Db 892 ACCAACTATTGAACCCACCGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 951
 QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 24417
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-08-846-762-1

Alignment Scores:

Pred. No.: 6,81e-108 Length: 24417
Score: 982.00 Matches: 190
Percent Similarity: 65.7% Conservativity: 70
Best Local Similarity: 48.0% Mismatches: 136
Query Match: 48.0% Indels: 0
DB: 2 Gaps: 0

US-10-673-786A-2 (1-396) x US-08-846-762-1 (1-24417)

```
QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 23597 CTGTTTCTGCGCTGAAATGACCGCGTGTATCCCATCTTGGCTGAACGACCTTC 23538
QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 23537 AACGCCGATACCCGTCGGCGAAGATCAACCTGGCGGTGGCGGTGACTACAAACGAGGAG 23478
QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 23477 GGGCGCATTCGCTGTCGTCGCGTGCAGGCTGCGAGAGCGCGCGCATCGAGCCAC 23418
QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 23417 GCGCGCGCGGTACTCGCATCGAAGGATCGCGCGCTACGACGAGCGGTACAGAAA 23358
QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 23357 CTGCTGTTTCGGTAAACGAGTCCGAGTGTGTCGCGCGCGCGGTGTCACGACCGCC 23298
QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 23297 GTCGCGCGCGCGCGCTCAAGCTCGCGCGCGGACTTCCTCAAGCGCTGTCGCCGAC 23238
QY 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 23237 GCCACCGTGGCCATCAGCGACCGGAGTGGGAAACACCGCGGCTGTTTCAAGCGGCC 23178
QY 141 GlyLeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 23177 GGCCTTCCCGGTGCAGAACTACCGCTATTACGACGCGCGCAATGGCGTGAACCGCGCC 23118
QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 23117 GGCCTGCTGGAAGACTGATGCCCTCGCGCGCGCGCTGATCGTGGTCTGCAACCGCTGC 23058
QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 23057 TGCCATAATCCGACCGCGGTGATCTCGAGCTGGACGACTGGAACAGGTGCTGGACGTG 22998
QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 22997 CTCAAGGCCAAGGGCGACGTGCGCTTCTCGACATCGCTACGAGGCTTCGGCAACGCG 22938
QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 22937 ATCGAGGAAGACCGCGCGGTGCGCTGTTCCGCCAGTGGCGGCTGAGCTTCTCGTT 22878
QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 22877 TCCAGCTGTTCTCCAAATCTTCTCGCTCTATGCGAAGCGGTGCGGCGCTCTCGATC 22818
QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 22817 GTGACCGAAAGCGCGATGAATCGCGCGCGTGTGCCAGGTGAAGCGGTGATCCGCG 22758
```

```
QY 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 22757 ACCAACTATTGAAACCCACCGACCGCGCGAGCGTCTCTCCTCGTGTCAACAGC 22698
QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 22697 CCGGAACATGCGGGGCCCTCTGGAGGAGGAGAACTGGGCGGAGATCGCGACCGCATCCGCGAC 22638
QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 22637 ATGCGCTGGCGATGTCGACCAACTGGCAGCCGCGCGCCAGCGGACTTCAGTTTC 22578
QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 22577 GTCGGTGCCAACAGTGGCATGTTCTCTATTCCGCGCTGACCGCGGACCGAGTCGAGCGC 22518
QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 22517 CTGAAGACCGAGTTCGGTATCTACGCCGTACGACCGCGCGGTATCTGCGTCCGCGCGCTG 22458
QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 22457 AACAAATCGAACCTGGAGACCATCACCAAGGCGCATCTGTCAGGTCTCTG 22410
```

RESULT 9

US-09-489-039A-4954/c
; Sequence 4954, Application US/09489039A
; Patent No. 6610836

GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4954
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4954

Alignment Scores:

Pred. No.: 3,45e-107 Length: 579
Score: 955.00 Matches: 179
Percent Similarity: 95.3% Conservativity: 5
Best Local Similarity: 92.7% Mismatches: 9
Query Match: 46.7% Indels: 0
DB: 3 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-489-039A-4954 (1-579)

```
QY 14 LeuGlyLeuAlaAspLeuPheArgAlaAspGlyArgProGlyLysIleAsnLeuGlyIle 33
Db 579 TTAGCTGGCGCATCTGTTTCGTCGCGATGACCGCCCTGAAAAATTAACCTCGGAATT 520
QY 34 GlyValTyrLysAspGluThrGlyLysThrProValLeuThrSerValLysLysAlaGlu 53
Db 519 GGTGTTTACAAGGATGAACCGGTAAAAACGCGCTTCTGACCGCGTCAAAAAACGACAG 460
QY 54 GlnTyrLeuLeuGluAsnGluThrThrLysAsnTyrLeuGlyIleAspGlyIleProGlu 73
Db 459 CAGTATCTGCTGGAAAAATGAAACCGACTATAAACATATCTGGGCGATCGATGGTATTCCTGAA 400
QY 74 PheGlyArgCysThrGlnGluLeuPheGlyLysGlySerAlaLeuIleAsnAspLys 93
Db 399 TTTGGTCTGCTCACCCAGGAGCTGCTGTTCCGTAAGCAATGCGATTATCGTGATAAA 340
QY 94 ArgAlaArgThrAlaGlnThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPhe 113
Db 339 CGCGCCGCGACCGCGCAGACACCGCGCGGTACCGGTGCGCTGCGGTGTCGCGCAGACTTC 280
```

```
Qy 114 LeuAlaLysAsnThrSerValLysArgValTrpValSerAsnProSerTrpProAsnHis 133
Db 279 CTCGCCAAACACCCAGCGTGAACGTGTGTGGTAACTCCAGAGCTGGCCGGAACCAT 220
Qy 134 LysSerValPheAsnSerAlaGlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGlu 153
Db 219 AANAGCGTATTCACTCTGCCGGGTGGAAGTGGCGAATACGCATATCTACGACGGCT 160
Qy 154 AsnHisThrLeuAspPheAspAlaLeuLeuAsnSerLeuAsnGluAlaGlnAlaGlyAsp 173
Db 159 AACACCGCGCTGGACTTTGATGGTCTGCTGGCCAGCTGAACGAAGCCAGCGGGCGAC 100
Qy 174 ValValLeuPheHisGlyCysCysHisAsnProThrGlyLeuAspProThrLeuGluGln 193
Db 99 GTGGTACTGTTCCAGCGGTGTCGCACCAACCGACCGGTATCGATCCGACGCTCGATCAG 40
Qy 194 TrpGlnThrLeuAlaGlnLeuSerValGluLysGlyTyr 206
Db 39 TGGCAGCAGCTGGCGCAGCTGTGCGTGGAAAAAGGCTGG 1
```

RESULT 10

US-09-252-991A-12968/c

; Sequence 12968, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 12968

; LENGTH: 1509

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-12968

Alignment Scores:

Pred. No.:	5,91e-101	Length:	1509
Score:	909.50	Matches:	185
Percent Similarity:	62.7%	Conservative:	62
Best Local Similarity:	47.0%	Mismatches:	146
Query Match:	44.5%	Indels:	1
DB:	3	Gaps:	1

US-10-673-786A-2 (1-396) x US-09-252-991A-12968 (1-1509)

```
Qy 2 PheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPheArg 21
Db 1290 TTCCGCCAAGTCCGCCGGTACCAGCGACCCGATCTCTGGCGCTCTCGACGCTACCGC 1231
Qy 22 AlaAspGluArgProGlyIleAsnLeuGlyIleGlyValTyrLysAspGluThrGly 41
Db 1230 AACGATCCGCGCGGACCAAGTACCTCGCGCGTGTCTACAGGATGCCCGAGGCG 1171
Qy 42 LysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGluThr 61
Db 1170 CTGACCCCGATCCTGCGCTCGGTGAACTCCCGAGCAGCGCTTGGTTCGACGAGAAACC 1111
Qy 62 ThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyA-gCysThrGlnGluLeu 81
Db 1110 ACCAAGACTAGTCCGCGGCGCACCGGATCGGTGTTCCGCGCGCTTGGCGGAACGT 1051
Qy 82 LeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaA-gThrAlaGlnThrPro 101
Db 1050 CGCTCGCGCGCGCTCGCCCTGTGCTGGAGCAACGCGCGCACGCCACCCAGACGCCC 991
```

```
Qy 102 GlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerValLys 121
Db 990 GCGCGCACCGCGCTTTCGCGCTGGCGGCGACTTCATCGCCCATTCCTGCCCGCGCGC 931
Qy 122 ArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAlaGly 141
Db 930 GGCATCTGGCTGAGCGACCGACCTGGCCGATCCACGAGACCCCTGTTCCGCGCGCGC 871
Qy 142 LeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAspAla 161
Db 870 CTGAAGGTTTCCCACTACCCCTACGTACGCGCGCCGACCAACCGC---CTGGATGTCCAGCG 814
Qy 162 LeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCys 181
Db 813 ATGCTTCTGCGCTGGAGCGCATTTCCCGAGGAGACGTGTGTGCTGCTGCATGCTGCTGC 754
Qy 182 HisAsnProThrGlyIleAspProThrLeuGluGlnTyrPheLeuAlaGlnLeuSer 201
Db 753 CACAAACCCGACCGGTTTCGACCTGAGCCACGACGACTGGCGAGGCTGCTCGAGCTGGTG 694
Qy 202 ValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGlyLeu 221
Db 693 CGTCGCGCAGCTGCTGCGCTGATCGACTTCGCTTACGAGGCTTTCGCGACGCTCTC 634
Qy 222 GluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMethHisLysGluLeuIleValAla 241
Db 633 GAGGAACACGCTGGCGCGTACGCTTTCGCGCGGAACTGCGCGAGGTGCTGCTGCTCAC 574
Qy 242 SerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeuVal 261
Db 573 AGTTCCTGCTGAAGAACTTCGCGCTGTACGCGCGCGCTGCGCGGCTGATGCTGCTGC 514
Qy 262 AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAla 281
Db 513 GCGCAGAACCCGAGAACTCCGACCTCGGTAGCCAACTGGCTTCTCGCCCGGAAAC 454
Qy 282 AsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsnAsp 301
Db 453 CTCTGCTGACCCCGCGCGCATGCGCGAGGTGCTCGCGCGATCTCTCGCGACAGC 394
Qy 302 AlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArgMet 321
Db 393 GAGTTGAAGGAGCTTTGGCAGGAAGGTGCAAGGCGATGCGCTCGCCCATCGCCAGCGCTG 334
Qy 322 ArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPheIle 341
Db 333 CGCATCGCGCTGCTCGAAGCCCTGGCGCGCCGCGCTGGCGCGCTTTCGCCCATGTC 274
Qy 342 IleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeu 361
Db 273 GCGCGCAACCGCGGATGTTTCTATACCGGACTGAGCGCGCGAGCGAGGTGCTGCGCTG 214
Qy 362 ArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
Db 213 CGGACGAGCAGCCCGCTTACCTGGTGTCCAGCGCGCGCGCCCAACGTCGCGGTATAGAC 154
Qy 382 ProAspAsnMetAlaProLeuCysGluAlaIleValAlaVal 395
Db 153 GCGCGCGCGCTCCACCGCTGGCGCAAGCATCGCCAGGTC 112
```

RESULT 11

US-09-252-991A-12507

; Sequence 12507, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

```

; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12507
; LENGTH: 1632
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12507

Alignment Scores:
Pred. No.: 6,7e-101 Length: 1632
Score: 909.50 Matches: 185
Percent Similarity: 62.7% Conservative: 62
Best Local Similarity: 47.0% Mismatches: 146
Query Match: 44.5% Indels: 1
DB: 3 Gaps: 1

US-10-673-786A-2 (1-396) x US-09-252-991A-12507 (1-1632)

QY 2 PheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPheArg 21
DB 442 TTCGCCAAGTCCGCCGCGTACCGCGGACCGGATCTCGGCTGCTCGACGCTACCGC 501
QY 22 AlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThrGly 41
DB 502 AACGATCCGCGCGGACAAAGCTGGACCTCGCGTGGTGTCTACAAAGGATGCCAGGCG 561
QY 42 LysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGluThr 61
DB 562 CTGACCCCGATCTCGGCTCGGTCGGAACCTCGCGAGCAGCGCTGGTCGAGCAGGAAC 621
QY 62 ThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGluLeu 81
DB 622 ACCAGAGTACGTCGCGCGCCACCGCGATCGCTGTCGCGCGCGCTGCGCGAATG 681
QY 82 LeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThrPro 101
DB 682 GCGCTCGCGCGCGCTCGCGCTGTTGCTGGAGCAACGCGCGACGCCACCCAGACGCG 741
QY 102 GlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerValLys 121
DB 742 GCGCGCACCGCGCGCTTGGCGCTGCGCGCGATCTCATCGCCCTGCGCGCGCGCG 801
QY 122 ArgValTyrValSerAsnProSerTyrProAsnHisLysSerValPheAsnSerAlaGly 141
DB 802 GCATCTGCTGAGCGACCGACCTGCGCGATCCAGACACCTGTTCCGCGCGCGCGCG 861
QY 142 LeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAspAla 161
DB 862 CTGAAGGTTTCCCACTACCCCTACGTCAGCGCGCGCAACCGC--CTGGATGTCGAGGCG 918
QY 162 LeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCys 181
DB 919 ATGCTGTGCTGCGCGCGCGCATTCGCCAGGAGAGCTGCTGCTGCTGCTGCTGCTG 978
QY 182 HisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeuSer 201
DB 979 CACAACCCGAGCGTTCGACCTGAGCCAGCAGCTGGCGGAGGAGGTCGACGCGGCG 1038
QY 202 ValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGlyLeu 221
DB 1039 CGTCGCGCGGAGCTGTCGCGCTGATCGATTCGCTACCGGCTTCGCGCGGCTGCTC 1098
QY 222 GluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleValAla 241
DB 1099 GAGGAAGACGCTGGCGCGGTACGCTGTTCCGCGCGGAACTGCGCGAGGTGCTGCTGCT 1158
QY 242 SerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValClyAlaCysThrLeuVal 261
DB 1159 AGTTCCTGCTCGAAGAACCTTCGCGCTGTATCCCGCGACCGCGTGGGCGCTGATCGCT 1218
QY 262 AlaAlaAspSerGluThrValAspAlaPheSerGlnMetLysAlaAlaIleArgAla 281

```

```

DB 1219 GCGGAGAACGCGGACGAGGCTACCGACCTGCGTAGCCAACTGCGCTTCCTCGCCGAAAC 1278
QY 282 AsnTyrSerAsnProAlaHisGlyAlaSerValAlaAlaThrIleLeuSerAsnAsp 301
DB 1279 CTCTGTCGACCCCGCGCGCATGCGCGAGGTGGTTCGCGCGCATCTTCGCGGACAGC 1338
QY 302 AlaLeuArgAlaIleTyrGluGlnLeuThrAspMetArgGlnArgIleGlnArgMet 321
DB 1339 GAGTTGAAGGAGCTTTGGCAGGAAGGTCGAAGGATCGCTCGCGCATCCGACGCTG 1398
QY 322 ArgGlnLeuPheValAsnThrLeuGlnGlyLysGlyAlaAsnArgAspPheSerPheIle 341
DB 1399 CGCATCGGCTGTCGAAGCCCTGCGCGCGGACGCGCTGCGCGAGGCTTCGCCCATGTC 1458
QY 342 IleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeu 361
DB 1459 GCGCGCAACGCGGCGATGTTTCTTATACCGGACTGAGCCCGCAGCAGGTGCTCGCTG 1518
QY 362 ArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
DB 1519 CGCAGCAGCAGCAGCGCTTTACCTGCTGTCAGCGCGCGGCAACGTCGCGGTATAGAC 1578
QY 382 ProAspAsnMetAlaProLeuCysGluAlaIleValAlaVal 395
DB 1579 GCGCGCGCGCTCGACCGCTGCGCGCAAGCCATCGCCCGAGGTC 1620

RESULT 12
US-09-543-681A-3805
; Sequence 3805, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3805
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-3805

Alignment Scores:
Pred. No.: 2,65e-101 Length: 702
Score: 908.00 Matches: 165
Percent Similarity: 86.0% Conservative: 25
Best Local Similarity: 74.7% Mismatches: 31
Query Match: 44.4% Indels: 0
DB: 3 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-543-681A-3805 (1-702)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
DB 37 ATGTTTGAGAAATCATTCGTCAGCAGCGCATCTTCTGGGTTTAGCTATGATTTTC 96
QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
DB 97 CGTTCGATCTCGTGAAACAAATTAACCTAGGATTTGGTGTATTAAGGATGAACA 156
QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
DB 157 GGTAAACACCTGTACTCACCACCGTTAAAAAAGCAGAAAAATATTTATTAGAAACGAA 216
QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
DB 217 TCCACATAAAATATCTTCCTATAGCGGTATTCCTGAGTTTGGTGGCGCTCACTCAAGCA 276
QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100

```


RESULT 14
US-09-540-236-315
; Sequence 315, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 315
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: M.catarrhalis
US-09-540-236-315

Alignment Scores:
Pred. No.: 1.85e-96 Length: 1209
Score: 871.50 Matches: 176
Percent Similarity: 59.7% Conservative: 61
Best Local Similarity: 44.3% Mismatches: 159
Query Match: 42.6% Indels: 1
DB: 3 Gaps: 1

US-10-673-786A-2 (1-396) x US-09-540-236-315 (1-1209)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 13 ATGTTTGAACATTAACCACTTATGCGAGCGGCCCAATTTTAGGCTTGATGATAAATAT 72
QY 21 ArgAlaAspGluArgProGly---LysIleAsnLeuGlyIleGlyValThrLysAspGlu 39
Db 73 GCCAAGACCCGACGACTGACATTAAGTCAATCTTGGTGGCGGTGACTATACCGAA 132
QY 40 ThrGlyLysThrProValLeuThrSerValLysLysAlaGluGlnTrpLeuLeuGluAsn 59
Db 133 GATGTCATTCCTGCTACTTGAATCGCTCAAAACCGCGATCCAAATCAACCAATCCA 192
QY 60 GluThrThrLysAsnTrpLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGln 79
Db 193 CCTCGTCCAGGAGGATTTGCAATGATGTTGCGCAGGTTATCGTAAGGCATGCCAA 252
QY 80 GluLeuLeuPheGlyLysSerAlaLeuIleAsnAspLysArgAlaArgThrAlaGln 99
Db 253 GATTTGTTATTTGGTAAATAATCATCTGCGAGTTTGGAGGGCGGTGTTGCCACGATTGCA 312
QY 100 ThrProGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSer 119
Db 313 ACTTTGGCGGTTCTGGTGGTAAAGTTGGTGTGATTTTATCCATGAGTGGTTCTCT 372
QY 120 ValLysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSer 139
Db 373 ATNGCAATGCTATGTGTCACCACTTGGCAATCATCTGGGTATTTTTCAGGGT 432
QY 140 AlaGlyLeuGluValArgGluTrpAlaTrpAspAlaGluAsnHisThrLeuAspPhe 159
Db 433 GCAGGTTTGGTGGCAAGATGCTTATATGATCCAAATACTTGGCGTAAATTT 492
QY 160 AspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGly 179
Db 493 GATGAGCTGTCGAATTTTCAAGACTTTAAATGAATAATGATGTAGTGTACTACACCA 552
QY 180 CysCysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGln 199
Db 553 TGTTCATTAACCACTGGTGGTGGATTTGACCCGTGACAGTGAATGATCTATTACAA 612
QY 200 LeuSerValGluLysGlyTrpLeuProLeuPheAspPheAlaTrpGlnGlyPheAlaArg 219
Db 613 ATCGTCAAGAGATAAAACCTTATTCATTTATGGATATTCCTCATCAAGTTTTCGGCAT 672

QY 220 GlyLeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIle 239
Db 673 GATATGGATGGTATGATCGTATGCCATTCGTGCGGTGGCGTGGATGATGACGGTATTT 732
QY 240 ValAlaSerSerTySerLysAsnPheGlyLeuTrpAsnGluArgValGlyAlaCysThr 259
Db 733 GTCTCTAATTCATTTTCAAAAACCTTGCTACCTTATGTTGAGCGGTGGCGGTTATCT 792
QY 260 LeuValAlaAlaAspSerGlyThrValAspArgAlaPheSerGlnMetLysAlaAlaIle 279
Db 793 GTCTGGTCCACCAACCAAGAGACAGATGTGGTACTTGGTCAATTAATAATTTACCGTC 852
QY 280 ArgAlaAsnTrpSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSer 299
Db 853 CGTCGATTTTATCAAGTCTGCACACGCAACTATGTTGTTGATCGCGTCATGAAT 912
QY 300 AsnAspAlaLeuArgAlaIleTrpGluGlnLeuThrAspMetArgGlnArgIleGln 319
Db 913 GATGAAGCATTTATTTGCGTATGATGATGATGATGATGATGATGATGATGATGATGAT 972
QY 320 ArgMetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyValaAsnArgAspPheSer 339
Db 973 GATATGCGCTCAAAAACCTCAAGACATTAAGCACTAACTCCCTGAGCGTATTTTACG 1032
QY 340 PheIleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeu 359
Db 1033 TATTTTACCAAAACAGCGTGTATGTTTACGACTTACAGGACTAACCAAGCAAGTGAT 1092
QY 360 ArgLeuArgGluGluPheGlyValTrpAlaValAlaSerGlyArgValaAsnValaGly 379
Db 1093 CGTTTGGCGTGAAGAAATTTGCGGTATATATGTTGAAATGCGCGTATGTTGCGAGGT 1152
QY 380 MetThrProAspAsnMetAlaProLeuCysGluAlaIleValaAlaValLeu 396
Db 1153 CTGAATAATTTCTATGATAGATATGTTGCCAGTGTATGCGTGAAGTGCTT 1203

RESULT 15
US-09-489-039A-1301
; Sequence 1301, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1301
; LENGTH: 1338
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1301

Alignment Scores:
Pred. No.: 4.39e-93 Length: 1338
Score: 844.50 Matches: 167
Percent Similarity: 60.7% Conservative: 74
Best Local Similarity: 42.1% Mismatches: 155
Query Match: 41.3% Indels: 1
DB: 3 Gaps: 1

US-10-673-786A-2 (1-396) x US-09-489-039A-1301 (1-1338)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 145 GTGTTTCAAAAGTTGACGCTACCGCGGACCCCTATCTTCTCTGATGAAGCGTTC 204
QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValThrLysAspGluThr 40
Db 205 AAAGAAGACCCGCGAAGCGACAAAGTCAACCTGAGTATCGGCTGCTACTACAACGATGAC 264

Search completed: March 15, 2006, 04:22:49
Job time : 3161 secs

Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeu---LeuGluAsn 59
Db 265 GGCATTATCCCGCAGCTGCAGCGCGTGGCGGAAGCGAAGCACGCCCTCAACCGCAGCGC 324
Qy 60 GluThrThrLysAsnTyrLeuGlyLeuAspGlyTyrProGluPheGlyArgCysThrGln 79
Db 325 CATGGCGCTCGCTGTATCTCCCAATGGAAGGGTTGACGGCTACCGCTCAGCGCATTCGC 384
Qy 80 GluLeuLeuPheGlyLysGlySerAlaLeuLeuAsnAspLysArgAlaArgThrAlaGln 99
Db 385 CCACCTGCTGTTGGCGCAGACATACCGCGCTTAAGCAAAATCGATTGCCTCTATTTCAG 444
Qy 100 ThrProGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSer 119
Db 445 ACGGTCCGCGGTTCGGGGCGGCTGAAGCTCGGGCGCCACTTCCTCAACAGTTACTTTCCT 504
Qy 120 ValLysArgValTyrValSerAsnProSerTyrProAsnHisLysSerValPheAsnSer 139
Db 505 GAGTCTCATGCTCGGTCAGGATCCGACCTGGGAAACCATCATCGCCATTTTTGAAGGG 564
Qy 140 AlaGlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPhe 159
Db 565 GCTGGCTTCGAAGTAAGTACTTACCCCTGGTTGATAAAGCACCAACGCGCTCGCTTT 624
Qy 160 AspAlaLeuLeuAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGly 179
Db 625 GAGGACCTGCTGGCGAGCGCTGCAACCGCTGCCGCGCCGCGATATGTGCTCTGCACCT 684
Qy 180 CysCysHisAsnProThrGlyLeuAspProThrLeuGluGlnTyrGlnThrLeuAlaGln 199
Db 685 TGCTGCACAAACCGACCGCGCGCAGCTGACGCCAGCCACAGTGGGACCGCGTGTGAG 744
Qy 200 LeuSerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArg 219
Db 745 GTATTGAAGCGCGCTCAGCTGATCCCTCTCTCGACATCGCTACCAAGGCGTTTGGCGGG 804
Qy 220 GlyLeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuLe 239
Db 805 GGGCTGGAAGAAGATGCTTACGCTATTTCGCGCCATCGCCAGCGCGGGATGCCGATGCTG 864
Qy 240 ValAlaSerGertySerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThr 259
Db 865 GTCAGCAACTCGTTCTCCAAATTTTCTCGCTGTACGGGGAGCGCTCGCGCGTCTGTCC 924
Qy 260 LeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaLe 279
Db 925 GTGCTGTGTGAAGACAGCGAACC CGCGGCGCGCTGTGTTGGGGCAGCTGAAGGCCACCGTG 984
Qy 280 ArgAlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSer 299
Db 985 CGCGCTAACTACTCCAGCCCGCAGCTTTGGCGCGCAGGTGTGTGGCGCAGCGTCTGAAC 1044
Qy 300 AsnAspAlaLeuArgAlaLeuTyrGluGlnLeuThrAspMetArgGlnArgIleGln 319
Db 1045 GATCCGGGCTGAAGAGCCAGCTGCGACGGCGGAAGTGGACGCCATTCGCGCGCACATCTCTG 1104
Qy 320 ArgMetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSer 339
Db 1105 ACCATGCGCCAGCGCTGGTTCGACGCGCTGCAGCAGGTTCGCCGGCAGCAAGTGGAT 1164
Qy 340 PheIleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeu 359
Db 1165 TACCTGTCAAGCAGCGCGGATGTTTACGCTATACCGGATTTAGCGCGCGCAGCGTGGAT 1224
Qy 360 ArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGly 379
Db 1225 CGTCTGGCGGATGATGTTGGCGTCTACTGATTCGACGCGACCATGTGCTGCGCGGG 1284
Qy 380 MetThrProAspAsnMetAlaProLeuCysGluAlaLeuValAlaValLeu 396
Db 1285 TTAACTCAGCTAATGTTTACGAGCTAGCAAGAGCTTTTGGCGCGGTAACTT3335
Db

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 15:51:10 ; Search time 1031 Seconds
(without alignments)
3176.211 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045

Sequence: 1 MFENITAPADPILGLDLF.....VAGMTPDNAPLCEAIVAVL 396

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q=/abs/ABSSWEB_spool/US10673786/runat_14032006_135621_22855/app_query.fasta_1
-DB=Published Applications NA Main -QFMT=fastap -SUFFIX=rnpsbm -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs805h
-USER=US10673786@CGN_1_1_1026@runat_14032006_135621_22855 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2045	100.0	1191	6	US-10-369-493-24472
2	2045	100.0	1191	7	US-10-673-786A-1
3	2045	100.0	14759	9	US-10-893-671-87
4	1654	80.9	1188	6	US-10-369-493-44812
5	1343	65.7	7977	9	US-10-795-159-536
6	1343	65.7	908766	9	US-10-795-159-685
7	1325	64.8	1830121	7	US-10-329-670-1
					Sequence 1, Appli

8	1325	64.8	1830121	8	US-10-158-865-1	Sequence 1, Appli
9	1325	64.8	1830121	9	US-10-981-687-1	Sequence 1, Appli
10	1318	64.4	966	6	US-10-369-493-23918	Sequence 23918, A
11	1311	64.1	1191	7	US-10-275-026A-187	Sequence 187, App
12	1310	64.1	17381	9	US-10-915-740A-40	Sequence 40, Appl
13	1310	64.1	2242716	9	US-10-915-740A-1068	Sequence 1068, Ap
14	1005.5	49.2	1185	6	US-10-369-493-35511	Sequence 35511, A
15	996	48.7	1188	6	US-10-369-493-31184	Sequence 31184, A
16	996	48.7	1194	6	US-10-369-493-28425	Sequence 28425, A
17	982	48.0	1191	6	US-10-369-493-32111	Sequence 32111, A
18	982	48.0	24417	6	US-10-216-209-1	Sequence 1, Appli
19	975	47.7	1191	6	US-10-369-493-39515	Sequence 39515, A
20	975	47.7	1191	6	US-10-369-493-39895	Sequence 39895, A
21	975	47.7	1194	6	US-10-369-493-39147	Sequence 39147, A
22	947	46.3	1194	6	US-10-369-493-37712	Sequence 37712, A
23	924	45.2	1185	6	US-10-369-493-33145	Sequence 33145, A
24	917	44.8	1185	6	US-10-369-493-32909	Sequence 32909, A
25	917	44.8	1481	7	US-10-425-114-35115	Sequence 35115, A
26	914	44.7	1489	7	US-10-425-114-20320	Sequence 20320, A
27	914	44.7	1817	7	US-10-767-701-14736	Sequence 14736, A
28	913	44.6	1456	7	US-10-425-114-32011	Sequence 22011, A
29	911.5	44.6	1185	6	US-10-369-493-37410	Sequence 37410, A
30	911	44.5	1508	7	US-10-425-114-31281	Sequence 31281, A
31	909	44.4	1203	6	US-10-369-493-41174	Sequence 41174, A
32	909	44.4	2731748	7	US-10-297-465A-1	Sequence 1, Appli
33	904.5	44.2	2152	8	US-10-425-115-12630	Sequence 12630, A
34	893	43.7	1188	6	US-10-369-493-32346	Sequence 32346, A
35	885	43.3	1775	7	US-10-424-599-96239	Sequence 96239, A
36	884.5	43.3	1188	6	US-10-369-493-44841	Sequence 44841, A
37	875.5	42.8	1188	6	US-10-369-493-23974	Sequence 23974, A
38	872.5	42.7	100848	7	US-10-672-787-39	Sequence 39, Appl
39	870	42.5	5976	7	US-10-437-963-84918	Sequence 84918, A
40	866	42.3	3913	8	US-10-425-115-103242	Sequence 103242, A
41	865	42.3	3913	8	US-10-425-115-179006	Sequence 179006, A
42	863	42.2	1932	7	US-10-424-599-29666	Sequence 29666, A
43	861	42.1	1817	7	US-10-767-701-14745	Sequence 14745, A
44	859	42.0	1690	7	US-10-425-114-33582	Sequence 33582, A
45	857	41.9	2084	8	US-10-425-115-179004	Sequence 179004, A

ALIGNMENTS

RESULT 1

US-10-369-493-24472
; Sequence 24472, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24472
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-369-493-24472

Alignment Scores:
Pred. No.: 6.32e-227 Length: 1191
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservativeness: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-369-493-24472 (1-1191)

QY 1 MetPheGluAenIleThrAlaAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Db 1 ATGTTTGGAGAACATTACCGCGCTCTCTGCCGACCGGATTTCTGGGCGCTGGCGGATCTGTTT 60

QY 21 ArgAlaAaspGluArgProGlyLyseAenLeuGlyIleGlyValTyrlsAaspGluThr 40
Db 61 CGTGCCGATGAACCTCCCGGCAAAATTAACCTCGGGATTTGGTGTCTATAAAGATGAGACG 120

QY 41 GlyLyseThrProValLeuThrSerValTyrlsLyseAlaGluGlnTyrlsLeuLeuAenGlu 60
Db 121 GGCAAAACCCCGGTACTGACCGCGTGAAGAAAGGCTGAAACAGTATCTGCTCGAAAAATGAA 180

QY 61 ThrThrLyseAenTyrlsLeuGlyIleAaspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 181 ACCACCAAAATTTACCTCGGCATTGACGGCATCCCTGGAATTTGGTTCGCTGCACCTCAGGAA 240

QY 81 LeuLeuPheGlyLyseGlySerAlaLeuIleAenAaspLyseArgAlaAargThrAlaGlnThr 100
Db 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCACGGCACAGACT 300

QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAaAaspPheLeuAlaLyseAenThrSerVal 120
Db 301 CCGGGGGGACCTACCGGCTACCGGCTACCGGCTACCGGCTACCGGCTACCGGCTACCGGCT 360

QY 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 361 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCCGAACCAATAAGAGCGTCTTTAACTCTGCA 420

QY 141 GlyLeuGluValArgGluTyrlsAlaTyrlsAaspAlaGluAenHisThrLeuAaspPheAasp 160
Db 421 GGCTCTGAAGAGTCTGTAATACGCTTATATGATCGGAAATACACACTCTTGACTTCGAT 480

QY 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCys 180
Db 481 GCATCTGATTAACACCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTCTGTTCCATGGCTGC 540

QY 181 CysHisAenProThrGlyIleAaspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 541 TGCCATAACCCCAACCGGTATCGACCCCTACGCTGGAACCAATGGCAAAACACTGGCACAACTC 600

QY 201 SerValGluLyseGlyTrpLeuProLeuPheAaspPheAlaTyrlsGlnGlyPheAlaAargGly 220
Db 601 TCCGTTTGAGAAAGGCTGGTACCCTGTTTACCTTACCGAGGCTTTGGCCCGGT 660

QY 221 LeuGluGluAaspAlaGluGlyLeuArgAlaPheAlaAlaMethHisLyseGluLeuIleVal 240
Db 661 CTGGAAGAGATGCTGAAGGACTCGCGCTTTCGCGCTATGTCATAAAGAGCTGATTTGT 720

QY 241 AlaSerSerTyrlsAenPheGlyLeuTyrlsAenGluArgValGlyAlaCysThrLeu 260
Db 721 GCCAGTTCTCTACTCTAAACCTTTGGGCTGTACACGAGGCTGTGGCGCTTGTACTCTG 780

QY 261 ValAlaAlaAaspSerGluThrValAaspArgAlaPheSerGlnMetLyseAlaAlaIleArg 280
Db 781 GTTCTGCGGACGTGAACCGTTTGTGTCGCGCATTCAGCCAAATGAAGCGGGGATTCGC 840

QY 281 AlaAenTyrlsSerAenProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
Db 841 GCTAACTACTCTAACCCACGACGACGCGCTTCTGTTGTCACCATCTCTGAGCAAC 900

QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAaspMetArgGlnArgIleGlnArg 320
Db 901 GATCGTTACGTGCGATTTGGGAAACAAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT 960

QY 321 MetArgGlnLeuPheValAenThrLeuGlnGluLyseGlyValAenAargAaspPheSerPhe 340
Db 961 ATGCGTCAGTTGTTGTCGTAATACCTGACGAAAGCGGCAAAACCCGCGACTTCAGCTTT 1020

QY 341 IleIleLyseGlnAenGlyMetPheSerPheSerGlyLeuThrLyseGluGlnValLeuArg 360

Db 1915 GCTAACTACTCTAACCCACAGACACGCGCTCTCTGTTGTGGCCACATCTCAGGAC 1856
Qy 301 AspAlaLeuArgAlaIleTrrpGluGlnGluLeuThrAspMetArgGlnArgGlnArg 320
Db 1855 GATGGTTACGTGGATTGGGACAGAGCTGACTGATATGCGCCAGCGTATTACGGT 1796
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLeuGlnGlnGlnGlnGlnGln 340
Db 1795 ATGGCTCAGTTGTTGCTCAATACGCTGAGGAAAGGCGGAAACCGCGACTTCAGCTTT 1736
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGlnGlnValLeuArg 360
Db 1735 ATCATCAACAGACGCGATGTTCTCTTCAGTGGCTTGACAAAGAACCAAGTCTCGGT 1676
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1675 CTGGCGAAGAGTTGGGTATATGCGGTGCTTCTGTCGCTAATATGTCGCGGATG 1616
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1615 ACACAGATAATGCTCCGCTGCTGCGAAGCGATTGTGGCAGTGTG 1568

RESULT 4

US-10-369-493-44812
; Sequence 44812, Application US/10369493
; Publication No. US20030233675A1

GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 44812
; LENGTH: 1188
; TYPE: DNA
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-44812

Alignment Scores:
Pred. No.: 1.71e-181 Length: 1188
Score: 1654.00 Matches: 310
Percent Similarity: 86.9% Conservative: 34
Best Local Similarity: 78.3% Mismatches: 52
Query Match: 80.9% Indels: 0
Gaps: 0

US-10-673-786A-2 (1-396) x US-10-369-493-44812 (1-1188)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTTGAAAATAATACAGCAGCAGCTCCGACCCCTATCTGTTAGCGGATAGCTTT 60
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 61 AAAGCTGATCTCGTGAAATAATAATCAATCTGGGTATCGGTGTTTACAAAGACGAACT 120
Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 121 GGTAACCCCTGCTCTGACCCAGCTTAAAGACGAGAAATTCCTCTCGTAACCGAA 180
Qy 61 ThrThrLysAsnTyrLeuGlyLysIleProGluPheGlyArgCysThrGlnGlu 80
Db 181 ACCACCAAAATATCTGGCAATAGTGGTACCTGAATTTGGCGGCTCACTCAGGA 240

Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 241 CTACTCTTCGGCAATACCAGCACTATCATCTGACAAACGCGCCGCTACTGTACAAGC 300
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 301 CAGGCGAACCAGTGCACCTGCGCAGATTTTATTGCTAAACAGCAATGCA 360
Qy 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 361 AAACGTCTCTGGATCAGTAACCAACATGGCCAAACATAAAGCGGTTTTCTTAGCGCT 420
Qy 141 GlyLeuGluValArgGluTyrAlaTyrAspAlaGluAenHisThrLeuAspPheAsp 160
Db 421 GGGTTAGAAATCCCGAGTACCAATATTCAATGCAGAAAGCATGCTCTGGATTTGAT 480
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 481 GGCATGTTGGCTAGCTGCTGAAGCTCAAGCAGGTGATGTTGTTCTTCACGGTTGC 540
Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 541 TGCATAAACCAACCGGTATCGATCCAAACGGCTGAACAGTGGCAGAAACCTGGCAGATTGA 600
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 601 TCTGCGGCAATGGCTGGCTGCTGTTGATTCTCCTTATCAGGTTTTTCCAGAGT 660
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaAlaMetHisLysGluLeuIleVal 240
Db 661 TTAGATGAAGATGCAGAGGCGCTGCGTATTTTACTAAACCATATAATGAATCGTT 720
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 721 GCCAGCTCTTACTCCAAAAATTTGGCTGTACAAATGAGCGTGTAGTGTGCTGCACCATC 780
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 781 GTTGCAACTGACAGCGATACCGCAGAAAAAGCATTCAGTCAGGCGAAATCTATTGTTCTGT 840
Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 841 ACTAACTACTCTAACCCACCGCACAGCGCTCTGTTGTTACTACAATTTCTGTCCAAT 900
Qy 301 AspAlaLeuArgAlaIleTrrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 901 GACGAATTTAAAGCAGAAATGGATTCAAGAACTGGCGACTATGCGTGAACGCAATTCGTCGT 960
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 961 ATGCGTCAGCTGTTTGTGAACACCTTCGAAAGAAAGAGTGCAGAAACAGACCTTTAGCTTT 1020
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1021 ATTATTAGCCAAATGCTATGTTTTCATTCACTGCTGCTGACTAAAGAAACAAAGTTGATCGT 1080
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1081 CTGCGTGAAGAATTCGGTATTTATTCGCGTCAAGTCTGCGGCGCATTAACGTAGCGGGCTTA 1140
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1141 ACACGTGAGAACATGGTTCCTCTGTGTGAAGCGGATTGTTGCAGTACTC 1188

RESULT 5

US-10-795-159-536
; Sequence 536, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: US/10/795,159

```

; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 536
; LENGTH: 7977
; TYPE: DNA
; ORGANISM: H. influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4108)..(4108)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7960)..(7960)
; OTHER INFORMATION: n = a, c, g, or t
US-10-795-159-536

Alignment Scores:
Pred. No.: 3,77e-144 Length: 7977
Score: 1343.00 Matches: 251
Percent Similarity: 78.3% Conservative: 59
Best Local Similarity: 63.4% Mismatches: 86
Query Match: 65.7% Indels: 0
Gaps: 0
DB:

US-10-673-786A-2 (1-396) x US-10-795-159-536 (1-7977)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1594 ATGTTGAACATATTAAAGCGGCACCGATCCCAATCTTAGGCTTAGGCGAAGCATTC 1653

Qy 21 ArgAlaAspGluAtpProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 1654 AAATCCGAACCTCGCGAAATAATAATTAATCTGGGTATTGGCGTTTATAAGATCGCGAA 1713

Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuGluAsnGlu 60
Db 1714 GGCACACCCCGCATTTATGCGTCAGTGAAGNAGCAGAAACGGCTTATTGTATGAAGAA 1773

Qy 61 ThrThrLysAsnTyrLeuGlyLysAspGlyLeuProGluPheGlyArgCysThrGlnGlu 80
Db 1774 AAAACGAAAAATTTATCTGACTATCGATCGTATTGCTGATTATTAACGAAACAAACGCG 1833

Qy 81 LeuLeuPheGlyLysGlySerAlaLeuLeuAsnLysArgAlaArgThrAlaGlnThr 100
Db 1834 TTACTTTTCGGTAAAGATTTCTGAAGTCATCAATCTAATCGTGCAGAACAGTACAAAGT 1893

Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 1894 TTAGCGGACACGCGCATTCAGCATTCGGCGAGATTTTAAACGTCACAACTAAAGCA 1953

Qy 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 1954 CAAATGTTTGATGAGCAGCCCACTTGGCCACACCAATGCCATTTTCAATGCCGTC 2013

Qy 141 GlyLeuGluValArgGluTyrAlaTyrThrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 2014 GGTATGACAAATTCGTGAATATCGTTATTATGATGCTGAACGCAACGCGCTTGAITGGAT 2073

Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 2074 CATTTACTTGAAGATTAAAGCAACAGCAGGAGCGGATGGTGGCTTTTACCGGTTGT 2133

Qy 181 CysHisAsnProThrGlyLeuAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 2134 TGCCATATCCGACTGGTATTGACCTTACTCCAGAACATGGCAAGATTAGCCGCACTT 2193

Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 2194 TCGGCGAAAAACGGTGGTGGCTGCTTTGACTTTGCTTATCAAGTTTAGCCCAATGGA 2253

```

```

Qy 221 LeuGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 2254 TTAGATGAAGATGCTTATGCTTGGCGCAAAACCAACAAAGAAATTTGTAGTA 2313

Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrArgGluValGlyValAcysThrLeu 260
Db 2314 CGGAGTTCAATCTCGAAAAACCTTTGGTTTATATGAACGTTTGGTGGCTTTACTCTT 2373

Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 2374 GTGGCAGAAATGCGAGAAATTCATCAACCGCATTAACCCAGTGAATCAATATTCTGT 2433

Qy 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 2434 ACACCTACTCTAAACCTCTCACTCAGCGCGCGCACTGTAGCAACAGTATTAAATGAC 2493

Qy 301 AspAlaLeuArgAlaIleTyrGluGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320
Db 2494 ACTCAACTTCGCCAAGAGTGGGAAATGAATTAATCTGAATGCGGCAACGATCAAAAAA 2553

Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnLysGlyValAlaAsnArgAspPheSerPhe 340
Db 2554 ATGGTCAATTTATTCGTTCACTTATTAAGAAATACGCTGCGGAAACAGATTTCACTTT 2613

Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 2614 ATCATTAACAAACGCGTATGTTCTCTTCACTGATTAATCTGGGGAACAGTGGATCGC 2673

Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 2674 TTAAGAAGAATTTGCCATTTATGCTTCTGCTGCTATCAATGTGCTGCTGATTT 2733

Qy 381 ThrProAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 2734 ACTGAAGATAATATTCGCTATTATGTAAGAACGATTTGGAAGTGCTT 2781

RESULT 6
US-10-795-159-685/c
; Sequence 685, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: US/10/795,159
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 685
; LENGTH: 908766
; TYPE: DNA
; ORGANISM: H. influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9223)..(9223)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (39640)..(39640)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (292404)..(292404)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (292495)..(292495)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (908604)..(908641)

```


OTHER INFORMATION: n is a, c, g, or t
US-10-795-159-685

Alignment Scores:

Pred. No.: 3,71e-141 Length: 908766
Score: 1343.00 Matches: 251
Percent Similarity: 78.3% Conservative: 59
Best Local Similarity: 63.4% Mismatches: 86
Query Match: 65.7% Indels: 0
DB: 9 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-795-159-685 (1-908766)

```
Qy 1 MetPheGluAenIleThrAlaAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Db 217833 ATGTTGAACATATTAAGCGGCACACGCCGATCCAATCTTAGGCTTAGGCGAAGCATTC 217774
Qy 21 ArgAlaAaspGluAatqProGlyLysIleAenLeuGlyIleGlyValTyLysAaspGluThr 40
Db 217773 AAATCCGAAACTCGCGAAATAAATAATTAACCTGGGTATTTGGCGTTTATAAAGATGCGCAA 217714
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyLeuLeuGluAasnGlu 60
Db 217713 GGCACAACCCGATATATCGTGCAGTGAAGAGCAGAAACCGCTTATTTGATAAGGAA 217654
Qy 61 ThrThrLysAenTyLeuGlyIleAaspGlyIleProGluPheGlyA-gCysThrGlnGlu 80
Db 217653 AAAACGAAAAATATCTGACTATCGATTCGATTCGATTAATAACGAAACAAACAAAGCG 217594
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAasnAspLysArgAlaAargThrAlaGlnThr 100
Db 217593 TTACTTTTCGGTAAGATCTCGAAGTCATCAAAATCTAATCGTGCAGAACAGTCAAAAGT 217534
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAaspPheLeuAlaLysAasnThrSerVal 120
Db 217533 TTAGCGGAAACAGCGCATTCAGCATTCGCGCAGAAATTTATAACGTCAAACTAAAGCA 217474
Qy 121 LysArgValTrpValSerAasnProSerTrpProAenHisLysSerValPheAasnSerAla 140
Db 217473 CAAAATGTTGGATCAGCACCCCAACTTGGCCAAACACCAATGCCATTTTCAATGCCGTC 217414
Qy 141 GlyLeuGluValArgGluTyAlaTyTrpAspAlaGluAenHisThrLeuAaspPheAasp 160
Db 217413 GGTATGACAATTCGTGAATATCGTATTATGATGCTGAACGCAAAAGCCCTTGTATGGGAT 217354
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCys 180
Db 217353 CATTTACTTGAGATTNAGCAGCAAGCAAGCGCATGGTGCTTTACACGGTTGT 217294
Qy 181 CysHisAenProThrGlyIleAaspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 217293 TGCCATAATCCGACTGGTATTGACCCCTACTCCAGAACAAATGGCAAGATTAGCCGCACTT 217234
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAaspPheAlaTyTrpGlnGlyPheAlaAargGly 220
Db 217233 TCGCGCAAAAACGGGTGGTCCACTCTTTGACTTTGCTTATCAAGGTTTAGCCAAATGGA 217174
Qy 221 LeuGluGluAaspAlaGluGlyLeuArgAlaPheAlaMetHisLysLeuLeuIleVal 240
Db 217173 TTAGATGAAGATGCTTATGGCTTCGGTCTTTTCGGCAACCAACCAAGAAATTTAGTA 217114
Qy 241 AlaSerSerTySerLysAasnPheGlyLeuTyTrpAenGluArgValGlyAlaCysThrLeu 260
Db 217113 GCGAGTTCAATCTCGAAAAAATTTGGTTTATATATGAACGTGTTGGTGCCTTTACTCTT 217054
Qy 261 ValAlaAlaAaspSerGluThrValAaspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 217053 GTGGCAGAAAAATGCAGAAATTCGATCAACCCGATTAACCCCAAGTGAATCAATATTATTCGT 216994
Qy 281 AlaAenTySerAasnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAasn 300
Db 216993 ACACCTACTACTACCCCTGCATCTTCACGGCGCGCAACTGTAGCAACAGTATTAAATGAC 216934
```

```
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAaspMetArgGlnArgGlnArg 320
Db 216933 ACTCAACTTCGCCCAAGAGTGGGAAATGAATTAACGAAATGCGCGAACGTATCAAAAAA 216874
Qy 321 MetArgGlnLeuPheValAenThrLeuGlnGluLysGlyAlaAenArgAaspPheSerPhe 340
Db 216873 ATGCGTCATTTATTCGTTTCAGTTATTAAAGAATAACGGTCGGGACAAAGATTTTCAGCTTT 216814
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 216813 ATCATTTGAACAAAAACGGTATGTTCTCTTTTCAGTGGATTAATCGGGAAACAAGTGGATCGC 216754
Qy 361 LeuArgGluGluPheGlyValTyAlaValAlaAaspGlyArgValAenValAlaGlyMet 380
Db 216753 TTAAGAAGAATAATTCOCATTTATGCTGTTCTGCTGCTATCAATGTGGCTGGTATT 216694
Qy 381 ThrProAaspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 216693 ACTGAAGATNATATTCGCTATTATTTATGTGAAGCATTTGTGAAGTGCTT 216646

RESULT 7
US-10-329-670-1
; Sequence 1, Application US/10329670
; Publication No. US20040018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; FILE REFERENCES: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
```

LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)

OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c

```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (152530)..(152530)

Alignment Scores:
Pred. No.: 1,27e-138 Length: 1830121
Score: 1325.00 Matches: 248
Percent Similarity: 77.8% Conservative: 60
Best Local Similarity: 62.6% Mismatches: 88
Query Match: 64.8% Indels: 0
DB: 7 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-329-670-1 (1-1830121)

Qy 1 MetPheGluAenThrAlaAlaProAlaAspProfileLeuGluLeuAlaAspLeuPhe 20
Db 1684147 ATGTTGAACATATCAAGCGGCACCGCATCCATCTTAGCTTAGGCGAAGCATTT 1684206
Qy 21 ArgAlaAspGluAargProGlyLysAlaLeuLeuGlyGlyValTyrLysAspGluThr 40
Db 1684207 AAATCCGAACTCCGAAATAAATCAATTTGGGTATTGGCGTTTATAAAGATGCGCAA 1684266
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 1684267 GGCACAAACCCCAATATATCAGCGGTAAAGAGCCGAAACCGATATTATTGTAAGAA 1684326
Qy 61 ThrThrLysAsnTyrLeuGlyLysAspGlyLysProGluPheGlyArgCysThrGlnGlu 80
Db 1684327 AAACCAAGAATATCTGACTATCGATGATGATGCGGATTTATACGACAAACAAAGCA 1684386
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuLeuLeuAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 1684387 CTCCTTTTCGGTAAGATTTCTGAATCTATCAATCTAATCGAGCAAGCAAGTACAAAGT 1684446
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 1684447 TTAGCGGAAACAGGTGCAATAGCATTCGCGCAATTTATTAACGCCAACTAAAGCA 1684506
Qy 121 LysArgValTyrProValSerAsnProSerThrProAsnHisLysSerValPheAsnSerAla 140
Db 1684507 CAAATGTTGGTGCAGCAGCACCAACTTGGCCAAACCAACATGCGATTTCAATGCTGTC 1684566
Qy 141 GlyLeuGluValAlaArgGluTyrAlaTyrThrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 1684567 GGTATGACCATCGTGAATATCGTATTATGATGCTGAACGCAAGCCCTTGATTGGGAA 1684626
Qy 161 AlaLeuLeuAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 1684627 CACTTATTAGAAATTTAAGCCCAAGCAAGCGCATGTTGGTGTTCACCGTTGT 1684686
Qy 181 CysHisAsnProThrGlyLysAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 1684687 TGCCATAATCGACTGGTATTGACCTACTCCAGACAAATGGCAAGAAATTAGCCGCACTT 1684746
Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 1684747 TCAGTAAATATGGTTGGTCCACTCTTGACTTTGCTTATCAAGGTTTAGCCACGGA 1684806
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuLeuVal 240
Db 1684807 TTAGATGAAGATGCTTATGGTTTACGTCTTTTCGCAACCAACCAAGAAATTTATTAGTG 1684866
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 1684867 GCGAGTTCATCTCGAAATACTTTGGTTTATATATGACGTGTTGGTGCACTTTTACCCTT 1684926
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaLeuArg 280
Db 1684927 GTGGCAGAAATTCAGAAATTTGCATCAACCTATTAACCCCAAGTAAATCAATTTATTCGC 1684986
Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrLeuLeuSerAsn 300
Db 1684987 ACCCTATCTACTAACCCAGCTTCTCACGGCGGGCGGACCGTACGACCAAGTATTAAATGAT 1685046

```

```

Qy 301 AspAlaLeuArgAlaLalleTrrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 1685047 GCTCAACTCCCAAGAAATGGGAAATGAATTAACCTGAAATGCGTGAACGATCAAAAA 1685106
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 1685107 ATGCGTCACTATTTCGTTTCAGTTATTAAAGAATATGTCGACAGAACAAAGATTTTCA 1685166
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1685167 ATCATTGAACAAACGTAATGTTTCTTTCAGTGGATTATACAGGGGAACAAGTGGATCGT 1685226
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1685227 TTAATAAATGAATTTGCCATTTACGCTTTCGTTCTGCTGATCAACGTTAGCTGGAATC 1685286
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1685287 ACAGAAGATAATATTCGCTATCTATGTGAAAGTATCTGTGAAGTACTT 1685334

RESULT 8
US-10-158-865-1
; Sequence 1, Application US/10158865
; Publication No. US20040203093A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:

```

NAME/KEY: misc_feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature

LOCATION: (107248)..(107248)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)

; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (152530)..(152530)

Alignment Scores:
Pred. No.: 1,27e-138 Length: 1830121
Score: 1325.00 Matches: 248
Percent Similarity: 72.8% Conservative: 60
Best Local Similarity: 62.6% Mismatches: 88
Query Match: 64.8% Indels: 0
DB: 8 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-158-865-1 (1-1830121)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1684147 ATGTTTGAACATATCAAGCGGACCGCATCCATCTTAGGCTTAGCGGAAGCAATTT 1684206
Qy 21 ArgAlaAspGluArgProGlyLyLeAsnLeuGlyLeGlyValTyLysAspGluThr 40
Db 1684207 AATCCGAAACTCGCGAAATAAATCAATTTGGGTATTTGGCGTTTATAAGATGCGCAA 1684266
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyLysLeuGluAsnGlu 60
Db 1684267 GGCACACCCCAATATGTCACGCGGTAAAGAGCCGAAACCGATTTATTGTATAGGAA 1684326
Qy 61 ThrThrLysAenTyLysGlyLeAspGlyLeProGluPheGlyArgCysThrGlnGlu 80
Db 1684327 AAACCCAGAAATATCTGACTATCGATCGATTCGCGATATTAACGAAACAAACAAAGCA 1684386
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 1684387 CTCCTTTCGTAAGATCTCGAAGTCATCAATCTAATCGAGCAGACAGTACAAAGT 1684446
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 1684447 TTAGCGGAAACAGGTGCAATACGCAATTCGCGCAAAATTTATTAACGCCAAACTAAAGCA 1684506
Qy 121 LysArgValTTPValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 1684507 CAATAATGTTGGWTCAGCACCAACTTGGCCAAACCAATGCGATTTTCAATGCCTGTC 1684566
Qy 141 GlyLeuGluValArgGluTyAlaTyTrpAspAlaGluAenHisThrLeuAspPheAsp 160
Db 1684567 GGTATGACCATTCGTGAATATCGTTATATGATCTGACCGCAAGCCCTTGATGGGAA 1684626
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 1684627 CACTTATTAGAAGATTTAAGCCAAAGCAAGCGCATGTGTGCTTTTACACGGTTGT 1684686
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGlnTrpGlnThrLeuAlaGlnLeu 200
Db 1684687 TGCCATAATCGACTGGTATTGACCCCTACTCCAGAAACAATGGCAAGAAATTAGCCGCACTT 1684746
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyGlnGlyPheAlaArgGly 220
Db 1684747 TCACCTAAAATGGTGTGGTGGCCACTCTTGACTTTGCTTATCAAGGTTTAGCCACCGGA 1684806
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 1684807 TTAGATGAAGATGCTTATGGTTTACGTGCTTTTCAGCAAAACCAACCAAGAAATTATTAGTG 1684866
Qy 241 AlaSerSerTySerLysAenPheGlyLeuTyrsnGluArgValGlyAlaCysThrLeu 260
Db 1684867 GCGAGTTTCATCTCCGAAAACTTTGGTTTATATATTAATGAACGTGTGTGGCATTTTACCCTT 1684926
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 1684927 GTGCGCAAAATGAGAAATGTCATCACTCATACCCCAAGTAAATCAATATTTCG 1684986
Qy 281 AlaAenTySerAenProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 1684986 GTGCGCAAAATGAGAAATGTCATCACTCATACCCCAAGTAAATCAATATTTCG 1684986

Db 1684987 ACCTATACTCTAAACCAGCTTCTCAGCGGGGCGGACCGTAGCAACAGATTTAAATGAT 1685046
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 1685047 GCTCAACTTCGCCCAAGAAATGGGAAATGAATTAACGTAAATGCGTGAACGCATCAAAAA 1685106
Qy 321 MetArgGlnLeuPheValAenThrLeuGlnGluLysGlyAlaAenArgAspPheSerPhe 340
Db 1685107 ATCGGTCACTTATTCGTTCAAGTTATTAAGAAATATGTCGCAAGAACCAAGATTTTCAGCTTT 1685166
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1685167 ATCATTTGAACAAAACCGTAATGTTTTCTTCAGTGGATTAAACAGGGGAAACAGTGGATCGT 1685226
Qy 361 LeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAenValAlaGlyMet 380
Db 1685227 TTAAGAAATGAATTTGCCATTTACGCTGTTCTGTTCTGCTATCAACGTAGCTGGAATC 1685286
Qy 381 ThrProAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1685287 ACAGAAGATAATATTCGCTATCTATGTGAAGTATCGTGAAGTACTT 1685334

RESULT 9
US-10-981-687-1
; Sequence 1, Application US/10981687
; Publication No. US20050131222A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2CID12
; CURRENT APPLICATION NUMBER: US 10/981,687
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US 10/158,865
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature

LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)

OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:

Alignment Scores:
Pred. No.: 1,27e-138 Length: 1830121
Score: 1325.00 Matches: 248
Percent Similarity: 77.8% Conservative: 60
Best Local Similarity: 62.6% Mismatches: 88
Query Match: 64.8% Indels: 0
DB: 9 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-981-687-1 (1-1830121)

Qy 1 MetPheGluAenIleThrAlaAapProAlaAapProIleLeuGlyLeuAlaAapLeuPhe 20
Db 1684147 ATGTTTGAACATATCAAGCGCCAGCCGATCCAATCTTAGGCTTAGCGGAAGCAATTT 1684206

Qy 21 ArgAlaAapGluAargProGlyLybIleAenLeuGlyIleGlyValTyLyAspGluThr 40
Db 1684207 AAATCCGAAACTCGGAAATAAATCAATTTGGGTATTCGGTTTATAAGATGCCAA 1684266

Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyLeuLeuGluAenGlu 60
Db 1684267 GGCACAAACCCCAATATGACCGGTAAAGAGCGCAAAACGATATTTGTATAAGAA 1684326

Qy 61 ThrThrLysAenTyLeuGlyIleAapGlyIleProGluPheGlyAargCysThrGlnGlu 80
Db 1684327 AAAACCAAGAAATATCTGACTATCGATGGTATTCGGGATTATAACGAAACAAACAGCA 1684386

Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAapLysAargAlaAargThrAlaGlnThr 100
Db 1684387 CTCCTTTTCGTAAGATTCCTGAAGTCATCAATCTAATCGAGCAAGACAGTACAAAGT 1684446

Qy 101 ProGlyGlyThrGlyAlaLeuAargValAlaAlaAapPheLeuAlaLysAenThrSerVal 120
Db 1684447 TTAGCGGAAACAGGTGATACGCATTCGGCAGCAATTTATTAAACGCCAACTAAAGCA 1684506

Qy 121 LysArgValTropValSerAenProSerTropAenHisLysSerValPheAenSerAla 140
Db 1684507 CAAAATGTTGGTWCAGCACCAACCAATTCGGCAACACCAATGCGATTTTCAATGCTGTC 1684566

Qy 141 GlyLeuGluValAargGluTyAlaTyTyAaspAlaGluAenHisThrLeuAapPheAasp 160
Db 1684567 GGTATGACCATTCGTGAATATCGTTATATGATCTGACCGCAAGCCCTTGATGGGAA 1684626

Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCys 180
Db 1684627 CACTATTAGAAGATTTAAGCAAGCAAGCAAGCGATGTGGTCTTTTACACGGTTGT 1684686

Qy 181 CysHisAenProThrGlyIleAaspProThrLeuGluGlnTropGlnThrLeuAlaGlnLeu 200
Db 1684687 TGCCATAATCCGACTGGTATGACCCCTACTCCAGAACAAATGGCAAGAAATTAGCCGCACTT 1684746

Qy 201 SerValGluLysGlyTropLeuProLeuPheAaspPheAlaTyGlnGlyPheAlaAargGly 220
Db 1684747 TCACCTAAAATGGTTGGTTGCCACTCTTTGACCTTTGCTTATCAAGGTTTAGCCCAACGGA 1684806

Qy 221 LeuGluAaspAlaGluGlyLeuAargAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 1684807 TTAGATGAAGATGCTTATGGTTTACGTGCTTTTCAGCAACCAACCAAGAAATATTAGTG 1684866

Qy 241 AlaSerSerTySerLysAenPheGlyLeuTyAenGluAargValGlyAlaCysThrLeu 260
Db 1684867 GCGAGTTTCATTCGAAAAAATTTGGTTTATATATATGAACGTGTGTGGTACATTTACCCCTT 1684926

Qy 261 ValAlaAlaAaspSerGluThrValAaspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 1684927 GTGCAGAAATGACAGAAATTCATCAACCTTATACCCCAAGTAAATCAATATTTCGC 1684986

Qy 281 AlaAenTySerAenProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
```

```
Db 1684987 ACCTTACTCTAACCCAGCTTCTCACGGGGCGGAGCTAGCAACAGTATTAAATGAT 1685046
Qy 301 AspAlaLeuAargAlaIleTropGluGlnGluLeuThrAaspMetArgGlnAargIleGlnAarg 320
Db 1685047 GCTCAACTTCGCCAAGAAATGGGAAATGAATTAACCTGAATCGTGAACGATCAAAAAA 1685106
Qy 321 MetArgGlnLeuPheValAenThrLeuGlnGluLysGlyAlaAenAargAaspPheSerPhe 340
Db 1685107 ATCGCTCACTTATTCGTTCAAGTTATTAAAAAGAAATATCGTGCAGAAACAAGATTTTCAGCTTT 1685166
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuAarg 360
Db 1685167 ATCATTTGAACAAACCGGTATGTTTCTTTCAGTGGATTAACAGGGGAACAGTGGATCGT 1685226
Qy 361 LeuAargGluGluPheGlyValTyAlaValAlaSerGlyAargValAenValAlaGlyMet 380
Db 1685227 TTAATAAATGAATTTGCCATTTACGCTGTTCTGCTCGTATCAACGATAGCTGGAATC 1685286
Qy 381 ThrProAaspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1685287 ACAGAAGATAATATTCGCTATCTATGTGAAAGTATCGTGAAGTACTT 1685334

RESULT 10
US-10-369-493-23918
; Sequence 23918, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23918
; LENGTH: 966
; TYPE: DNA
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-23918

Alignment Scores:
Pred. No.: 1,4e-142 Length: 966
Score: 1318.00 Matches: 246
Percent Similarity: 86.3% Conservative: 32
Best Local Similarity: 76.4% Mismatches: 44
Query Match: 64.4% Indels: 0
DB: 6 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-369-493-23918 (1-966)

Qy 75 GlyArgCysThrGlnGluLeuLeuPheGlyLysGlySerAlaLeuIleAenAaspLysArg 94
Db 1 GGCGGGGTGATTGAGNACTGCTTTTGGAAAGTACCAAGTCCCATCGTACTGAAAAACGT 60

Qy 95 AlaAargThrAlaGlnThrProGlyGlyThrGlyAlaLeuAargValAlaAaspPheLeu 114
Db 61 GGCGGTACCGCTACAAAGCCAGCGGAAACAGGCGCACTGCGGTACTGCGCTGATTTTATT 120

Qy 115 AlaIleAenThrSerValLysAargValTropValSerAenProSerTropAenHisLys 134
Db 121 GCCAAGCAAACTAATGCCAAACCGTTTGGATCAGTAACCCCAACCTGGCCAAACCATAAA 180

Qy 135 SerValPheAenSerAlaGlyLeuGluValAargGluTyAlaTyTyAaspAlaGluAen 154
Db 181 GGTGTTTTCAGCGCAGCGTTTAGAGATCCGCGAATATACTATTACGATGCAGAAAAA 240
```


Qy 302 AlaLeuArgAlaIleTrrpGluGlnGluLeuThrAspMetArgGlnAArgileGlnArgMet 321
 Db 907 GATTTGAAAGCACAAATGATTCGCGAACTCGATGAAATTCGCGCGCGTATCAAAAGCCATG 966
 Qy 322 ArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPheIle 341
 Db 967 CGCCAAAATTTGTCGAGTTGCTCAAGCCAAAGGTGCAACTCAAGACTTTTGATTTTCATT 1026
 Qy 342 IleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGlnGlnValLeuArgLeu 361
 Db 1027 ATCGACAAACGCGAGTTCTCTTCAGCGGCTTGACTCCCGAACAGTCGACCGGTTTA 1086
 Qy 362 ArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
 Db 1087 AAAACGAGTTTGCCATTATGCGCTCGCTCCGCGCGCATCAACGTCGCGCGCATTACC 1146
 Qy 382 ProAspAsnMetAlaProLeuGlnGluAlaIleValAlaValLeu 396
 Db 1147 GACGACACATCGATTATCTGTGCGAAAGTATCGTNAAGTACTG 1191

RESULT 12

US-10-915-740A-40
 ; Sequence 40, Application US/10915740A
 ; Publication No. US20050191316A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Frazer, Claire M.
 ; APPLICANT: Hickey, Erin
 ; APPLICANT: Peterson, Jeremy
 ; APPLICANT: Tettelin, Hervé
 ; APPLICANT: Venter, J. Craig
 ; APPLICANT: Masignani, Vega
 ; APPLICANT: Galeotti, Cesira
 ; APPLICANT: Mora, Manroza
 ; APPLICANT: Ratti, Giulio
 ; APPLICANT: Scarselli, Maria
 ; APPLICANT: Scarlato, Vincenzo
 ; APPLICANT: Rappuoli, Rino
 ; APPLICANT: Pizzi, Mariagrazia
 ; APPLICANT: Grandi, Guido
 ; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
 ; FILE REFERENCE: 002441.00090
 ; CURRENT APPLICATION NUMBER: US/10/915,740A
 ; PRIOR FILING DATE: 2004-08-11
 ; PRIOR APPLICATION NUMBER: 09/806,866
 ; PRIOR FILING DATE: 1999-10-08
 ; PRIOR APPLICATION NUMBER: USN 60/103,794
 ; PRIOR FILING DATE: 1998-10-09
 ; PRIOR APPLICATION NUMBER: USN 60/132,068
 ; PRIOR FILING DATE: 1999-04-30
 ; PRIOR APPLICATION NUMBER: PCT/US99/25373
 ; PRIOR FILING DATE: 1999-10-08
 ; NUMBER OF SEQ ID NOS: 1068
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 40
 ; LENGTH: 17381
 ; TYPE: DNA
 ; ORGANISM: Neisseria meningitidis
 US-10-915-740A-40

Alignment Scores:
 Pred. No.: 8,01e-140 Length: 17381
 Score: 1310.00 Matches: 244
 Percent Similarity: 78.0% Conservative: 64
 Best Local Similarity: 61.8% Mismatches: 87
 Query Match: 64.1% Indels: 0
 DB: 9 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-915-740A-40 (1-17381)

Qy 2 PheGluAsnIleThrAlaIleProAlaAspProfileLeuGlyLeuAlaAspLeuPheArg 21
 Db 8597 TTCNAGCACATCGAAGCGCGCCCGCATCCGATTCTCGGTTTGGCGAAGCGTTCAAA 8656

Qy 22 AlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThrGly 41
 Db 8657 GCCGAAACCCGCCCGCCGCAAGTCAACCTCGCATCGCGCTTTATATAAGAGCGATCCGCC 8716
 Qy 42 LysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGluThr 61
 Db 8717 GCACACCCCTCGTCAAGCGCGTCAAGAAAGCGGAAACCGCTGTGTGGAAAGCGAAACC 8776
 Qy 62 ThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGluLeu 81
 Db 8777 ACCAAAACACTACCTACCATCGACGCGGTTCGCGACTACAAACGCGCAACCCAAATCCTG 8836
 Qy 82 LeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaThrGlnAlaGlnThrPro 101
 Db 8837 CTGTTCCGCAACAGCACGAAATCATCGCCAGCGCTCGCGCCCAAAACAGCGCAAGCCTC 8896
 Qy 102 GlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerValLys 121
 Db 8897 GCGCGTACGCGCGCATTCGCGTATTGCGGCGCGAGTTTGCCAAACGCCAGTTGAACGCGCAA 8956
 Qy 122 ArgValTrrpValSerAsnProSerTrrpProAsnHisLysSerValPheAsnSerAlaGly 141
 Db 8957 ACCATCTGGATTCCAAATCGACTTCGCCCAACACACACGCCATCGCCAAAGCGGTCTGT 9016
 Qy 142 LeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAspAla 161
 Db 9017 ATCAAGACAAACCTTATCGTTACTATGATGCGCCCAACACACGTTTGGATTGGAGCGGC 9076
 Qy 162 LeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCys 181
 Db 9077 ATGATTGAGGACTTGAGCCAAAGCGCAAAAGGCGACATCGCTCTGTCGCGGTCTGTCG 9136
 Qy 182 HisAsnProThrGlyIleAspProThrLeuGluGlnTrrpGlnThrLeuAlaGlnLeuSer 201
 Db 9137 CACAATCCTACCGGCATCGACCTACGCCCAACAAATGGGAAACTTTGGCAAACTTTCT 9196
 Qy 202 ValGluLysGlyTrrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGlyLeu 221
 Db 9197 GCCGAAAGCGTGGTTGCGCTGTTGACTTTGCTTACCAAGGCTTCGCGCAATGGTTTG 9256
 Qy 222 GluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMethHisLysGluLeuIleValAla 241
 Db 9257 GAAGAAGATGCTTACCGCTTCGCGGTTCCTTGAACACAAATACAGAAATTCGATGTCG 9316
 Qy 242 SerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeuVal 261
 Db 9317 AGCTCTTATTCAAAACCTTCGATATGTACACAGCGCGTTCGCGCGGTTCACTTTGGTG 9376
 Qy 262 AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAla 281
 Db 9377 GCCGAAGATGAAGAAACAGCAGCGCCGCGCCACAGGCAAGTCAAAACCATCATCCGTACC 9436
 Qy 282 AsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsnAsp 301
 Db 9437 TTGTATTCCAACCCGCTTCACACGCGTGCAACACCATTCGCGTGTGTGTGAAATAATGAT 9496
 Qy 302 AlaLeuArgAlaIleTrrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArgMet 321
 Db 9497 GATTTGAAGCACAAATGGATTGCGGAACCTCGAATGCGCGCGCCGCGCATCAAGCCATG 9556
 Qy 322 ArgGlnLeuPheValAsnThrLeuGlnGlyGlyAlaAsnArgAspPheSerPheIle 341
 Db 9557 CGCCAAAATTTTTCGCGGTTCCTCAAGCGCAAGGTGCAAGCCAAACCTTTGATTTTCATT 9616
 Qy 342 IleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGlnGlnValLeuArgLeu 361
 Db 9617 ATCAAAACAAACGCGTATGTTCTTTTCAGCGGCTTGACTCCCGCAAGTTCGACCGCGCTG 9676
 Qy 362 ArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
 Db 9677 AAAAACGAGTTTGCCATTATGCGCTCGCGCGCGCATCAACGTCGCGCGCATTACC 9736

Qy 382 ProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 9737 GACAACAACATCGATTATCTGTGCGAAGCATCGTGAAGTACTG 9781

RESULT 13

US-10-915-740A-1068/C

; Sequence 1068, Application US/10915740A

; Publication No. US20050191316A1

; GENERAL INFORMATION:

; APPLICANT: Frazer, Claire M.

; APPLICANT: Hickey, Erin

; APPLICANT: Peterson, Jeremy

; APPLICANT: Tettelin, Hervé

; APPLICANT: Venter, J. Craig

; APPLICANT: Massignani, Vega

; APPLICANT: Galeotti, Cesira

; APPLICANT: Mora, Manroza

; APPLICANT: Ratti, Giulio

; APPLICANT: Scarselli, Maria

; APPLICANT: Scarlatto, Vincenzo

; APPLICANT: Rappuoli, Rino

; APPLICANT: Pizza, Mariagratia

; APPLICANT: Grandi, Guido

; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use

; FILE REFERENCE: 002441.00090

; CURRENT APPLICATION NUMBER: US/10/915,740A

; CURRENT FILING DATE: 2004-08-11

; PRIOR APPLICATION NUMBER: 09/806,866

; PRIOR FILING DATE: 1999-10-08

; PRIOR APPLICATION NUMBER: USN 60/103,794

; PRIOR FILING DATE: 1998-10-09

; PRIOR APPLICATION NUMBER: USN 60/132,068

; PRIOR FILING DATE: 1999-04-30

; PRIOR APPLICATION NUMBER: PCT/US99/25373

; PRIOR FILING DATE: 1999-10-08

; NUMBER OF SEQ ID NOS: 1068

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1068

; LENGTH: 2242716

; TYPE: DNA

; ORGANISM: Neisseria meningitidis

US-10-915-740A-1068

Alignment Scores:

Pred. No.:	9,448-137	Length:	2242716
Score:	1310.00	Matches:	244
Percent Similarity:	78.0%	Conservative:	64
Best Local Similarity:	61.8%	Mismatches:	87
Query Match:	64.1%	Indels:	0
DB:	9	Gaps:	0

US-10-673-786A-2 (1-396) x US-10-915-740A-1068 (1-2242716)

Qy 2 PheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPheArg 21
Db 562972 TTCAGACATCGAAGCGCCCGCCGATTCCTCGTTTGGCGAAGCGTTCAA 562913
Qy 22 AlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThrGly 41
Db 562912 GCCGAAACCCCGCCGAAAGATCAACCTCGGCATCGCGGTTTATAAAGACGCGATCCGCG 562853
Qy 42 LysThrProValThrSerValLysAlaGluGlnTyLysLeuLeuGluAsnGluThr 61
Db 562852 CGGACACCCCTCGTCAAGACCGTCAAGAGACCGGAAACCGCTGTGGAAGCGAAACC 562793
Qy 62 ThrLysAsnTyLysLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGluLeu 81
Db 562792 ACCAAACTACCTACCATCGACGCGTTCGCGACTACACGCGCAACCCCAATCTCG 562733
Qy 82 LeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThrPro 101
Db 562732 CTGTTCGCAAGACACCAATCATCGCAGCCGTCGCGCAAAACAGCGCAAGCCTC 562673

Qy 102 GlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerValLys 121
Db 562672 GCGCGTACGGCGCATTCGTTATTCGCGCGAGTTTGCACAAACGCGAGTTGAACGCGCAA 562613
Qy 122 ArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAlaGly 141
Db 562612 ACCATCTGGATTTCATCCAGCTTGGCCCAACCAACGCGCATCGCAAGCGGTCTCGT 562553
Qy 142 LeuGluValArgGluTyAlaTyTrpAspAlaGluAsnHisThrLeuAspPheAspAla 161
Db 562552 ATCCAGACAAACCTTATCTATGATGCGCCCAACACGCGTTTGGATTGGGACGCG 562493
Qy 162 LeuLeAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCys 181
Db 562492 ATGATTGAGGACTTGAAGCAAGCGCAAAAGCGCACATCGTCTCTGCACGCGTCTGC 562433
Qy 182 HisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeuSer 201
Db 562432 CACAATCTTACCGGCATCGACCTACGCCCAACCAATGGAAACCTTTGGCAAACTTTCT 562373
Qy 202 ValGluLysGlyTrpLeuProLeuPheAspPheAlaTyTrpGlnGlyPheAlaArgGlyLeu 221
Db 562372 GCCGAAAGAGCTGGTTGCCGCTGTTGACTTTGCCTACCAAGGCTTCGCAATGGTTG 562313
Qy 222 GluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuValAla 241
Db 562312 GAAGAAGATGCTACGGCTTCGCGCTGTTCTTGAACAACAATACAGAATTCGTGATGCC 562253
Qy 242 SerSerTyLysAsnPheGlyLeuTyTrpAsnGluArgValGlyAlaCysThrLeuVal 261
Db 562252 AGCTCTTATTCCAAAACCTTCGTTATGACCAACGAGCGCGTTCGCGCGCTTCATTTG 562193
Qy 262 AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAla 281
Db 562192 GCCGAAAGATGAAGAACAGAGCGCGCCGACAGCAAGTCAAAACCATCATCGTACC 562133
Qy 282 AsnTyLysAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsnAsp 301
Db 562132 TTGTATTCCAACCGCGCTTCACACGCTGCGAACCACCATTCGCTGCTGTGAAAAATGAT 562073
Qy 302 AlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArgMet 321
Db 562072 GATTTGAAGCACAAATGGAATTCGCGAACTCGATGAAATGCGCGCGCATCAAAAGCCATG 562013
Qy 322 ArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPheIle 341
Db 562012 CGCAAAATTTGCGGGTTCCTCAAGCCAAAGTGCAGCCAAACTTTGATTTCAAT 561953
Qy 342 IleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeu 361
Db 561952 ATCAACAAACCGGTATGTTCTCTTTTCAGCGCTTGACTCCCGAACAAGTCGACGCGCTG 561893
Qy 362 ArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
Db 561892 AAAACGAGTTTGCATTTATCGCTCCGCTCCGCGCGCATCAACGTCGCGCGCATTACC 561833
Qy 382 ProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 561832 GACAACAACATCGATTATCTGTGCGAAGCATCGTGAAGTACTG 561788

RESULT 14

US-10-369-493-35511

; Sequence 35511, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

```
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 35511
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Mesorhizobium loti
US-10-369-493-35511

Alignment Scores:
Pred. No.: 3.89e-106 Length: 1185
Score: 1005.50 Matches: 199
Percent Similarity: 65.9% Conservatives: 62
Best Local Similarity: 50.3% Mismatches: 132
Query Match: 49.2% Indels: 3
DB: 6 Gaps: 1

US-10-673-786A-2 (1-396) x US-10-369-493-35511 (1-1185)
Qy 1 MetPheGluAenIleThrAlaAProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Db 1 ATGTTCCAGACCTCGAGCCAGCTCCGCCGACAAAGATCTTGGCCCTCATCGGCTCTAT 60
Qy 21 ArgAlaAaspGluArgProGlyLysIleAenLeuGlyLeuGlyValTyrLysAaspGluThr 40
Db 61 CGCGCCGATCCGCTCCCAACAGAGTCAGCTCGCGCTCGGCTGTACAAAGGATCGCGAC 120
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 121 GGCAAGACCCAGAGTATCGCGCTCGCGGCGACACCGGCTTCAACGGCTGCTGAACAGCCAG 180
Qy 61 ThrThrLysAenTyrLeuGlyLysAlaAaspGlyProGluPheGlyArgCysThrGlnGlu 80
Db 181 GACACCAAGACCTATCTCGGCTCGCGGCGACACCGGCTTCAACGGCTGATGGCCAG 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuLeuAenAaspLysArgAlaArgThrAlaGlnThr 100
Db 241 CTCGCTTCGGCCCGCGCGGACATG-----ACGGCAATTCGGCGGCCAGCGC 291
Qy 101 ProGlyThrGlyAlaLeuArgValAlaAaspPheLeuAlaLysAenThrSerVal 120
Db 292 CCTGGCGGCTCGGCTCGCTTGGCTGCGGAGCTGCTCAAGCGGACCCGCTTCGGAC 351
Qy 121 LysArgValTyrValSerAenProSerTyrProAenHisLysSerValPheAenSerAla 140
Db 352 GCCACGATCTGGCTGGACCCAACTTTCGACGCGGCTTCCGCTGATCGCGCTGCC 411
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAaspAlaGluAenHisThrLeuAaspPheAasp 160
Db 412 GGCCTGCAGATCCGCAATATCCCTATTTCGACGCGGCTTCCGCTGCTGCTGCC 471
Qy 161 AlaLeuLeuAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCys 180
Db 472 GACATGTGCGCGGCTTTCGAGACCGCAAGAGCGGCGAGCTGTGCTGCTGCAAGCGGTCG 531
Qy 181 CysHisAenProThrGlyLysLeuAenProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 532 TGCCACAAATCAACCGGCGCAATCTGATGCGCGGAGTGGGAGCGGCTCACCGATCTG 591
Qy 201 SerValGluLysGlyTyrLeuProLeuPheAaspPheAlaTyrGlnGlyPheAlaAargGly 220
Db 592 ATGCTGAGCGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651
Qy 221 LeuGluGluAaspAlaGluGlyLeuArgAlaPheAlaAmetHisLysGluLeuLeuVal 240
Db 652 CTCGAGCGCGAGCTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 711
Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyAlaCysThrLeu 260
Db 712 GCCTCGAGCTGCTCGAAGAAATTTCCCGCTTACCGCGACCGCTGTTCGGCGGACGATGGTT 771

RESULT 15
US-10-369-493-31184
; Sequence 31184, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 31184
; LENGTH: 1188
; TYPE: DNA
; ORGANISM: Burkholderia cepacia
US-10-369-493-31184

Alignment Scores:
Pred. No.: 4.95e-105 Length: 1188
Score: 996.00 Matches: 191
Percent Similarity: 65.2% Conservatives: 67
Best Local Similarity: 48.2% Mismatches: 138
Query Match: 48.7% Indels: 0
DB: 6 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-369-493-31184 (1-1188)
Qy 1 MetPheGluAenIleThrAlaAProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Db 1 CTGTCTTCGCGCGCTCGAATCTTCTCCCGGACCCGATTCCTGGGCTTGAACGAAGCCTTC 60
Qy 21 ArgAlaAaspGluArgProGlyLysIleAenLeuGlyLeuGlyValTyrLysAaspGluThr 40
Db 61 AATGCGGATACGCGCGACCAACCAAGGTCAACCTTGGCGTTCGCGCTGCTGCTGCTGCTGCTG 120
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 121 GGCAAGATTCCTGCTGCGCGCTTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
```

Search completed: March 15, 2006, 18:35:53
Job time : 4484 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 16:06:55 ; Search time 442 Seconds
(without alignments)
2089.040 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045

Sequence: 1 MFENITAPADPIGLADLF.....VAGWTPDNMPLCAIVAVL 396

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 8023312 seqs, 1165852854 residues

Total number of hits satisfying chosen parameters: 16046624

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q=/abs/ABSSWEB_spool/US10673786/runat_14032006_135624_22906/app_query.fasta_1
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=rnpbnp -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss03h
-USER=US10673786 @CGN_1_1_431 @runat_14032006_135624_22906 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:

1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
5: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
11: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
12: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2011	98.3	1191	12	US-11-114-922-75
2	1317	64.4	1269	8	US-10-467-657-6535
3	842	41.2	1191	8	US-10-467-657-1059
4	831.5	40.7	1194	8	US-10-979-821-31
					Sequence 75, Appl
					Sequence 6535, Ap
					Sequence 1059, Ap
					Sequence 31, Appl

5	831.5	40.7	1194	12	US-11-114-922-31	Sequence 31, Appl
6	818	40.0	1260	8	US-10-979-821-5	Sequence 5, Appl
7	818	40.0	1260	12	US-11-114-922-5	Sequence 5, Appl
8	815	39.9	1170	8	US-10-979-821-1	Sequence 1, Appl
9	815	39.9	1170	12	US-11-114-922-1	Sequence 1, Appl
10	810	39.6	1260	8	US-10-979-821-3	Sequence 3, Appl
11	810	39.6	1260	12	US-11-114-922-3	Sequence 3, Appl
12	781	38.2	2343	12	US-11-136-527-2278	Sequence 2278, Ap
13	727.5	35.6	1251	7	US-10-932-182A-82350	Sequence 82350, A
14	727.5	35.6	1251	7	US-10-932-182A-82350	Sequence 82350, A
15	721.5	35.3	1389	7	US-10-932-182A-4671	Sequence 4671, Ap
16	721.5	35.3	1389	7	US-10-932-182A-4671	Sequence 4671, Ap
17	632.5	30.9	1239	8	US-10-979-821-7	Sequence 7, Appl
18	632.5	30.9	1239	12	US-11-114-922-7	Sequence 7, Appl
19	616.5	30.1	1356	7	US-10-932-182A-1541	Sequence 1541, Ap
20	616.5	30.1	1356	7	US-10-932-182A-1541	Sequence 1541, Ap
21	597.5	29.2	1356	7	US-10-932-182A-77511	Sequence 77511, A
22	597.5	29.2	1356	7	US-10-932-182A-77511	Sequence 77511, A
23	348	17.0	560	6	US-09-925-065A-479173	Sequence 479173, A
24	348	17.0	560	6	US-09-925-065A-479173	Sequence 479173, A
25	281	13.7	1400	12	US-11-136-527-6374	Sequence 6374, Ap
26	191	9.3	594	6	US-09-925-065A-739968	Sequence 739968, A
27	153.5	7.5	1290	8	US-10-793-626-3173	Sequence 3173, Ap
28	153.5	7.5	2883	8	US-10-793-626-3173	Sequence 4366, Ap
c 29	152.5	7.5	2631	8	US-10-793-626-3760	Sequence 3760, Ap
30	124.5	6.1	1230	12	US-11-055-822-115	Sequence 115, App
31	124.5	6.1	1230	12	US-11-055-822-485	Sequence 485, App
32	119	5.8	1119	7	US-10-932-182A-3622	Sequence 3622, Ap
33	119	5.8	1119	7	US-10-932-182A-3622	Sequence 3622, Ap
c 34	112.5	5.5	922	8	US-10-750-185-38121	Sequence 38121, A
c 35	112.5	5.5	922	8	US-10-750-623-38121	Sequence 38121, A
36	111	5.4	1356	8	US-10-467-657-1849	Sequence 1849, Ap
37	109.5	5.4	1161	7	US-10-932-182A-3002	Sequence 3002, Ap
38	109.5	5.4	1161	7	US-10-932-182A-3002	Sequence 3002, Ap
39	108	5.3	2529	12	US-11-136-527-3251	Sequence 3251, Ap
40	105.5	5.2	1194	9	US-11-232-405A-41	Sequence 41, Appl
c 41	103.5	5.1	4235	8	US-10-793-626-4041	Sequence 4041, Ap
42	103	5.0	5190	12	US-11-091-883-31	Sequence 31, Appl
43	103	5.0	5190	12	US-11-091-883-177	Sequence 177, Appl
44	99	4.8	1182	8	US-10-979-821-9	Sequence 9, Appl
45	99	4.8	1182	12	US-11-114-922-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-11-114-922-75
; Sequence 75, Application US/1114922
; Publication No. US20050282260A1
; GENERAL INFORMATION:
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; CURRENT FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 75
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
US-11-114-922-75

Alignment Scores: 2.76e-209 Length: 1191
Pred. No.: 2011.00 Matches: 390
Score: 99.2% Conservative: 3
Percent Similarity:


```
Best Local Similarity: 98.5%      Mismatches: 3
Query Match: 98.3%              Indels: 0
DB: 12                          Gaps: 0

US-10-673-786A-2 (1-396) x US-11-114-922-75 (1-1191)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTTTGAGAACATTACCGCGCTCTGCCGACCGGATCTGGGCGTGGCGATCTGTTT 60
Qy 21 ArgAlaAspGluArgProGlyIleAenLeuGlyIleGlyValTyrLysAspGluThr 40
Db 61 CGTCCGATGAACCTCCGCGGAAATTAACCTCGGATTTGGTCTCTATTACGATGAGACG 120
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 121 GGCAAAATCCCGGTACTGACCGCGCTGNAAGAGCTGNAACAGTATCTGCTCGAAATGAA 180
Qy 61 ThrThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 181 ACCACCAAACTTACCTCGGCATTCAGCGCATCCCTGAATTTGGTCTGCACTCAGAA 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaAargThrAlaGlnThr 100
Db 241 CTGCTGTTTGGTAAGGTAGCGCCCTGATCAATGACAAACGTGTCGACGCGCACACT 300
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 301 CCGGGGGCTCTGGCGCAGTACGCGTGGCTGCCGATTTCTGGCAAAATATACAGCGTT 360
Qy 121 LysArgValTrpValSerAsnProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 361 AAGCGTGTGTGGTGAGCAACCCCAAGCTGGCGGAACCATAGAGCGTCTTTAACTTGCA 420
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTrpAlaGluAenHisThrLeuAspPheAsp 160
Db 421 GGTCTGGAAGTTCTGTAATAGCTTATTATGATCGGAAATCACTCTTTGACTTCGAT 480
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 481 GCATGTATTACAGCTGATGAAGCTCAGCTGGCGACGAGTGTGCTGTTCCATGGCTGC 540
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 541 TGCCATAACCAACCGGTATCGACCTAGCTGGAACAAATGGCAACACATGGCACAACTC 600
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 601 TCCGTTGAGAAAGGTGTTTACCGCTGTTTGACTTCGCTTACCAAGGGTTTTGCCCGTGT 660
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
Db 661 CTGGAAGAAGATGCTGAAGAGCTGCGCGCTTTCGCGCTATGCAATAAGAGCTGATTGTT 720
Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyAlaCysThrLeu 260
Db 721 GCCAGTTCCTACTCTAATAAATTTGGCCCTGTACAAACGAGCGTGTGGCGCTTGACTCTG 780
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 781 GTTGTCGCGACAGTGAACCGCTTGATCGCGCATTTAGCCCAATGAAGCGCGGATTCGC 840
Qy 281 AlaAenTyrSerAenProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
Db 841 GCTAACTACTCTTCCCCACGACGACGCGCTCTCTGTTGTTCACCATCTGAGCAAC 900
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgGileGlnArg 320
Db 901 GATCGGTACGTGGGATTTGGGAACAAGCTGACTGATATATGCGCAGCGTATTACGGT 960
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnLysGlyAlaAenArgAspPheSerPhe 340
Db 961 ATGCGTCAAGTTGTTCTCAATACGCTGCGGAAAGCGCAAAACCGCGACTTCAGCTTT 1020
```

```
341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
1021 ATCATCAACAGAACGCGCATGTTCTCTTTCAGTGGCTGACAAAAGAACAGTCTCGT 1080
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyValValAenValAlaGlyMet 380
Db 1081 CTGCGGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTCGCTAAATGTGGCCGGGATG 1140
Qy 381 ThrProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1141 ACACCAAGTAACATGCTCCGCTGTGCGAAGCGATTGTGCGAGTGCTG 1188

RESULT 2
US-10-467-657-6535
; Sequence 6535, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCES:
; CURRENT APPLICATION NUMBER: US/10/467,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6535
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6535

Alignment Scores:
Pred. No.: 1,63e-133      Length: 1269
Score: 1317.00           Matches: 244
Percent Similarity: 77.7%      Mismatches: 62
Best Local Similarity: 61.9%      Indels: 88
Query Match: 64.4%           Gaps: 0
DB: 8

US-10-673-786A-2 (1-396) x US-10-467-657-6535 (1-1269)

Qy 2 PheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPheArg 21
Db 88 TTCAAGCACATCGAAGCGCGCCCGCGATCCGATTCGCTGGCGGAGCGTTCANA 147
Qy 22 AlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValTyrLysAspGluThrGly 41
Db 148 GCCGAACCCCGCCCGGAAAGTCACTCGGATCGCGCTGTACAAAGACGATCCGGC 207
Qy 42 LysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGluThr 61
Db 208 CGCACACCCATTTGTCAAAGCGCTGAAAGAGCTGNAAGAGCGCTGTGGAAGCGAAACC 267
Qy 62 ThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGluLeu 81
Db 268 ACCAAAAATTTACCTGACATCGAGCGCTGCGCGATACCAACGAGCAACCAATCTCTG 327
Qy 82 LeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaAargThrAlaGlnThrPro 101
Db 328 CTGTTTGGCAAGACCAAGAAATCATCCGCGCGCGCGCAAAACAGCGAAGCGCTT 387
Qy 102 GlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerValLys 121
Db 388 GCGGTACAGCGCATTTGCTGATTTGCGCGCAATTTGCCAAACGCTCAGTTGAACGCGCAA 447
Qy 122 ArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAlaGly 141
Db 1411 ATGCGTCAAGTTGTTCTCAATACGCTGCGGAAAGCGCAAAACCGCGACTTCAGCTTT
```

```

Db 448 ACCATCTGGATTCCAAATCCGACCTGGCCCAAAACCAACGCGCATTTGCCAAACGGGTGCGT 507
Qy 142 LeuGluValArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAspPheAspAla 161
Db 508 ATCCAGACAAACCTTATCGCTACTATGATGCGCCAAACACTGCTTGGATTGGACGGC 567
Qy 162 LeuLeuAenSerLeuAenGluAlaGlnAlaGlyAspValValPheHisGlyCysCys 181
Db 568 ATGATTGAAGATTGAACCAAGCCAAAGGCGACATCGTCTGCTGACGGCTGCTGC 627
Qy 182 HisAenProThrGlyLeuAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeuSer 201
Db 628 CACAACCCCTACCGGTATCGACCTACGCCCGCAACCAATGGAAACCTTTAGCAAACTTCT 687
Qy 202 ValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGlyLeu 221
Db 688 GCCGAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 747
Qy 222 GluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuLeuValAla 241
Db 748 GAAGAAGATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 807
Qy 242 SerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyAlaCysThrLeuVal 261
Db 808 AGCTCTTATTCAAAACCTTGGCGATGATACACGAGCGTGTGCGTGCATTCATCTTGGTG 867
Qy 262 AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaAlaAla 281
Db 868 GCCGAAGATGAAGAAACAGCAGCCCGCCGACCAACCAATCAACCATCATCTGCTACC 927
Qy 282 AsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsnAsp 301
Db 928 TTGTATTCCAAACCGCTTACATGCTGCGACACACCATTCGCTGCTGTGTGAATAATGAT 987
Qy 302 AlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArgMet 321
Db 988 CATTTGAAGCACAATGATGCTGCGAACTTGACGAATGCGCGCGCATCAAGCCCATG 1047
Qy 322 ArgGlnLeuPheValAsnThrLeuGlnLysGlyAlaAenArgAspPheSerPheIle 341
Db 1048 CCCCAAAATTTGTCAGTTGCTCAAGCCAAAGTTCAACCAAGACTTTCATTTATTTAT 1107
Qy 342 IleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeu 361
Db 1108 ATTGAACAAACGGTATGCTCTCTTTCAGCGCTTGTACTCCCGAACAAAGTTGACCGTTG 1167
Qy 362 ArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
Db 1168 AAAACGAGTTTGCATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1227
Qy 382 ProAspAenMetAlaProLeuCysGluAlaIleValAlaVal 395
Db 1228 GACGACAAACATCGATTATCTGTGTGAAGCATCGTGAAGTA 1269

```

RESULT 3

```

; Sequence 1059, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Rita
; APPLICANT: PIZZAZ Maria Rita
; APPLICANT: PASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1059

```

```

; LENGTH: 1191.
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1059

Alignment Scores:
Pred. No.: 1.03e-81 Length: 1191
Score: 842.00 Matches: 165
Percent Similarity: 58.8% Conservative: 68
Best Local Similarity: 41.7% Mismatches: 163
Query Match: 41.2% Indels: 0
DB: 8 Gaps: 0

US-10-673-786a-2 (1-396) x US-10-467-657-1059 (1-1191)

Qy 1 MetPheGluAenHisThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTACACACATCATGTAATCTATCCCGCGACCCGATTTTGGTTTGGTTCGAAACCTTC 60
Qy 21 ArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValTyrLysAspGluThr 40
Db 61 AAAACGACCCCGCGCGCCGCAAAAGTCAATTTGACATAGGCAATTTATTTTCGACGACGAA 120
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 121 GCGAGAATGCGCGTATTTGGAATCCGTAAGCGCGTGCAGAAACCCGCGCGCGCGCA 180
Qy 61 ThrThrLysAenTyrLeuGlyLysIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 181 GCGCGCTCGCGCTACTCCGATGAGGGTTTGGACGTTTACCGAGCGCGGTGCGAGCAT 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
Db 241 TTGTGTTTGGCAAGGCAACCCCGCTCGCGAGGAGCGCATCGTTACCGTACAGACT 300
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 301 CTGGCGCGCTCGCGCGCTCAAGTCTGAGCGCGGATTTCTGCGACCGCTGTTTCCGAA 360
Qy 121 LysArgValTyrValSerAsnProSerTyrProAenHisLysSerValPheAenSerAla 140
Db 361 GCGCGCGCTACGTCAGCGATCGGATTTGGGACCAACCATCGCGCGCATTTTGAAGCGCA 420
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAspPheAsp 160
Db 421 GGTTCGAGGTGCGCACTTACCATATTTACGACCTGCGCACCTGTCGGCGTGAATTCGAC 480
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 481 GAAATGACCGGCTTTTTCACACCCCTGCGCCGCAAAACAGCGTCTCATCTCCATCCCTGC 540
Qy 181 CysHisAenProThrGlyLysAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 541 TGCCACAAACCCGCGCGGTGATGTCGGAACCGCAATGGGACGAAGTGTGTGACAGATC 600
Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 601 ATCAAAACCGCAAACTGATTCGTTTATGACATCGCTACCAAGGCTTCGCGCGCGAT 660
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 661 TTGGACACGATGCTACGCGCTCGCAAAAGCGGTGGAATGGATTTGCGCGCTGTTGTC 720
Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyAlaCysThrLeu 260
Db 721 AGCAATTCGTTTCAAAAACCTGTCGCTTACGCGCAACCGCTCGCGCGGTTCAGCGGTG 780
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 781 GTTTCGCCCAATAAAGAAAGACCGGATTTGGTTGCGGACCACTCAATTTCCCGTCCGC 840
Qy 281 AlaAenTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300

```

Db 841 CGCATCTACTCCAGCCGCCGCCGCGTATATATCGCCGCCGAGTGATGAACAGC 900
Qy 301 AspAlaLeuArgAlaIleTrrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 901 TCCGAACCTTACGCTTGTGCGCAACGAAAGTTATATATGATGCGCGCATCCGTCG 960
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 961 ATGGCGCAGAACTTACGCGGTATTGACGCCCGGAATCCCGACCGCGATTTCACTTAT 1020
Qy 341 IleIleLysGlnAsnGlyMetPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1021 TTCAATCAACACGCGCGAATGTTGCGCTACACGGGGTTGACGTGGGCAAGTCCGCGG 1080
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1081 TTGGCGCAGCAATTTGCGCTTACTCTGCGGATCCGCGAGGATGCGCTCGCGGCTG 1140
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1141 AATACATCAAAATATCACTTATGTCGCGATGCGCTTGCAGAAAGTGTG 1188

RESULT 4

US-10-979-821-31

; Sequence 31, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.
; APPLICANT: HICKS, PAULA W.
; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; APPLICANT: WEINER, DAVID P.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; FILE REFERENCE: PRODUCTION OF MONATIN AND ITS PRECURSORS
; CURRENT APPLICATION NUMBER: US/10/979,821
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 31
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Escherichia coli

US-10-979-821-31

Alignment Scores:
Pred. No.: 1.45e-80 Length: 1194
Score: 831.50 Matches: 167
Percent Similarity: 60.7% Conservative: 74
Best Local Similarity: 42.1% Mismatches: 155
Query Match: 40.7% Indels: 1
DB: 8 Gaps: 1

US-10-673-786A-2 (1-396) x US-10-979-821-31 (1-1194)

Qy 1 MetPheGluAsnIleThrAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 GTGTTTCAAAAGTTGAGCTTACGCTGGCGACCCGATTTCTACGCTATGAGCGGTTT 60
Qy 21 ArgAlaAspGluArgProGlyIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 61 AAAGAAGACCCCTCGCAGCAGCAAAAGTGAATTAAGTATCGGTCTGTACTACAACGAAGAC 120
Qy 41 GlyLysThrProValLeuThrSerValLysLysGluGlnTyrLeu---LeuGluAsn 59
Db 121 GGAATTATTCACAACTGCAAGCCGTGGCGAGCGGAAGCGCCCTGAATGCGCAGCCT 180

Qy 60 GluThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGln 79
Db 181 CATGGCGCTTCGCTTTTATTTCATCCGATGGAAGGGCTTAACCTGCTATGCCATGCCATTCG 240
Qy 80 GluLeuLeuPheGlyGlySerAlaLeuLeuAsnAspLysArgAlaArgThrAlaGln 99
Db 241 CCCTGCTGTTTGGTGGGACCATCCGGTACTGAAACACAGCGCGTGAGCAACCATTCAA 300
Qy 100 ThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSer 119
Db 301 ACCCTGGCGGCTCCGGGCAATGAAAGTGGCGGATTTCTGAAACGCTACTTCCCG 360
Qy 120 ValLysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSer 139
Db 361 GAATCAGCGCTCTGGTTCAGCGATCTACCTCGGAAACCGCGTAGCAATATTCCGCGGG 420
Qy 140 AlaGlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAsnHisThrLeuAspPhe 159
Db 421 GCTGGATTCGAAGTGAGTACTTACCCTGGTATGACGAACGCGACTAACGGCGTCGGCTT 480
Qy 160 AspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGly 179
Db 481 AATGACCTGTGGCGAGCTGAAACATTAACCTGCCCGCAGTATTGTGTTGCTGCATCCA 540
Qy 180 CysCysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGln 199
Db 541 TGTTCACCAACCAACCGGTCGATCTCACTAATGATCAGTGGGATGCGGTGATTGAA 600
Qy 200 LeuSerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArg 219
Db 601 ATTCTCAAAGCCCGCAGCTTATTCCTATTCCTGATATTCCTATCAAGGATTTGGTGCC 660
Qy 220 GlyLeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIle 239
Db 661 GGTATGGAAGAGGATGCTACGCTATTCGCCCATTCGCCAGCGCTGATACCGCTCTG 720
Qy 240 ValAlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThr 259
Db 721 GTGAGCAATTCGTTCTCGAAATTTTCTCCCTTTACGGCGAGCGCGTGGCGGACTTCT 780
Qy 260 LeuValAlaAspSerGluThrValAspAlaPheSerGlnMetLysAlaAlaIle 279
Db 781 GTTATGTGAAGATCCGAAGCCGCTGCGCGCTACTGGGCGCAATTTGAAAGCAACAGTT 840
Qy 280 ArgAlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSer 299
Db 841 CGCGCAACTACTCTCAGCCCGCGAATTTGGTGGCAGGTGGTGGCTGCAGTCTGAAT 900
Qy 300 AsnAspAlaLeuArgAlaIleTrrpGluGlnGluLeuThrAspMetArgGlnArgIleGln 319
Db 901 GACGAGCATTTGAAGCCAGCTGGCTGGCGGAAGTAGAGAGATCGCTACTCGCATTTCTG 960
Qy 320 ArgMetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSer 339
Db 961 GCAATGCGTCAGGAATTTGGTGAAGGATTTAAGCACAGAGATGCCAGAACCAATTCGAT 1020
Qy 340 PheIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeu 359
Db 1021 TATCTGCTTAATCAGCGCGCATGTTCAAGTATACCGGTTTAAAGTGGCGCTCAGGTGAC 1080
Qy 360 ArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGly 379
Db 1081 CGACTACGTCAAGAAATTTGGTGTCTATCTCATCCAGCGCGTGCATGTGTGTCGCGGG 1140
Qy 380 MetThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1141 TTAATACCGCAAAATGTACAACTGTGTGCAAGCGCTTCTGCTGCGGTGATG 1191

RESULT 5

US-11-114-922-31

; Sequence 31, Application US/11114922

; Publication No. US20050282260A1

DB:	8	Gaps:	5
US-10-673-786A-2 (1-396) x US-10-979-821-5 (1-1260)			
Qy	1	MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe	20
Db	76	ATGTGACCGCCCTGGAAGCCGACCGCGGACCAAGATCTCTGCAACTGATCCAGATGTTTC	135
Qy	21	ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr	40
Db	136	CGCGAGGATGCGCGCGGACNAGATCGATCTGGCGTGGCGTCTACAGAGCCCGACC	195
Qy	41	GlyLysThrProValLeuThrSerValLysLeuGluInTyrLeuLeuGluAsnGlu	60
Db	196	GGGCTCACCCCGGTATCGCGCGGTGAAGCCCGCGGAGCGCGGTCTCGGAGGTGCG	255
Qy	61	ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu	80
Db	256	ACCACCAAGACCTACACCGGCTTGGCGGCGCGCGGAGCGCGGTCTCGGAGGTGCG	315
Qy	81	LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr	100
Db	316	CTGATCTCTC-----GCAGGCGCGGTCCCGCGGACCGCGGTGGCGTCTCGGTCGCCACC	366
Qy	101	ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal	120
Db	367	CCCGGGCGACCGCGCGGTGGCGTCTAGGCGCTCGAGCTGATCCGATGGCGTCCGCCGAG	426
Qy	121	LysArgValTrpValSerAsnProSerTyrProAsnHisLysSerValPheAsnSerAla	140
Db	427	GCCACTGTCTGGATCTCGAAGCCCGGACCTGGCGGACCACTCTGCGATCGTGAATATCTC	486
Qy	161	AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys	180
Db	547	GGCTTGATGGAGGATCTGGCCAGGTGAAGCGCGGCGACGCTGGTGTCTGTCACCGGTGC	606
Qy	181	CysHisAsnProThrGlyIleAspProThrLeuGluInTyrGlnThrLeuAlaGlnLeu	200
Db	607	TGCCACAAACCGACCGCGCGCAACCCGACCGGTGCGTGGCGTCTCGGAGAGC	666
Qy	201	SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly	220
Db	667	CTGGCCCGGACAGCGCGGTGGCGCTGATCGACCTCGCTTATCGAGGCTTTCGGGACGGG	726
Qy	221	LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal	240
Db	727	CTCGAGATGGATGGCGCGGCGACGCGCTTCTGGCCACAGACTCGCCGAGGTGCTGATC	786
Qy	241	AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu	260
Db	787	CGGCGCTCTGCTCGAAGAACTTCGGCATCTACCGGAGCGAAGCGGATCTCTGATCGCC	846
Qy	261	Val-----AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAla	278
Db	847	ATCGGCGAGGCGCGCGCGGCGGACCGGTGAG-----GCCAACCTCAACTTCTCTG	897
Qy	279	IleArgAlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeu	298
Db	898	AACCGGACAGAACTACTCTTCGCGCGGACCATGGCGCGGCTCGTGACCATGATCTCTC	957
Qy	299	SerAsnAspAlaLeuArgAlaIleTyrGluGlnLeuThrAspMetArgGlnArgIle	318
Db	958	GAGGACGAGACGCTGAGCGCGGACTGGAAGCGGAGCTCGAGAGGTGGCGCTCAACATG	1017
Qy	319	GlnArgMetArgGlnLeuPheValAsnThrLeuGln--GluLysGlyAlaAsnArgAsp	337
Db	1018	CTGACGCTGGCGCCCGACGCTTGGCGGTGGCGGCGGAGACCGGCTCGAACCGC---	1074
Qy	338	PheSerPheIleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGln	357

Db	1075	TTCCGGCTTCGTGCCGAGCATCGCGCATGTTCTCGCGCTTCGGGATCACGCCGCCGAG	1134
Qy	358	ValLeuArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnVal	377
Db	1135	GTGAGCGGCTGGCGGACCGGACGCGGTCTACATGTTGGCGGCTCGCGGCTGAACATC	1194
Qy	378	AlaGlyMetThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu	396
Db	1195	GCGGCGGTGAACACCGGACCGCTGCTGGCGCGCGGTGGCGCAAGGTGCTG	1251
RESULT 7			
US-11-114-922-5			
; Sequence 5, Application US/11114922			
; Publication No. US20050282260A1			
; GENERAL INFORMATION:			
; APPLICANT: HICKS, PAULA M.			
; APPLICANT: MCFARLAN, SARA C.			
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE			
; FILE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS			
; FILE REFERENCE: 023829-0396			
; CURRENT APPLICATION NUMBER: US/11/114,922			
; CURRENT FILING DATE: 2005-04-26			
; PRIOR APPLICATION NUMBER: 10/422,366			
; PRIOR FILING DATE: 2003-04-23			
; PRIOR APPLICATION NUMBER: 60/374,831			
; PRIOR FILING DATE: 2002-04-23			
; NUMBER OF SEQ ID NOS: 91			
; SOFTWARE: PatentIn Ver. 3.3			
; SEQ ID NO 5			
; LENGTH: 1260			
; TYPE: DNA			
; ORGANISM: Rhodobacter sphaeroides			
US-11-114-922-5			
Alignment Scores:			
Pred. No.:	4,61e-79	Length:	1260
Score:	818.00	Matches:	171
Percent Similarity:	58.1%	Conservative:	61
Best Local Similarity:	42.9%	Mismatches:	157
Query Match:	40.0%	Indels:	10
DB:	12	Gaps:	5
US-10-673-786A-2 (1-396) x US-11-114-922-5 (1-1260)			
Qy	1	MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe	20
Db	76	ATGCTGACCGCCCTGGAAGCCGCGGACCGCGGACCAAGATCTCTGCAACTGATCCAGATGTTTC	135
Qy	21	ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr	40
Db	136	CGCGAGGATGCGCGCGGACNAGATCGATCTGGCGTGGCGTCTACAAAGGACCGGACC	195
Qy	41	GlyLysThrProValLeuThrSerValLysLeuGluInTyrLeuLeuGluAsnGlu	60
Db	196	GGGCTCACCCCGGTATCGCGCGGTGAAGCCCGCGGAGCGCGGTCTCGGAGGTGCG	255
Qy	61	ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu	80
Db	256	ACCACCAAGACCTACACCGGCTTGGCGGCGCGGAGCGCGGTCTCGGAGGTGCG	315
Qy	81	LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr	100
Db	316	CTGATCTCTC-----GCAGGCGCGGTCCCGCGGACCGCGGTGGCGTCTCGGTCGCCACC	366
Qy	101	ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal	120
Db	367	CCCGGGCGACCGCGCGGTGGCGTCTAGGCGCTCGAGCTGATCCGATGGCGTCCGCCGAG	426
Qy	121	LysArgValTrpValSerAsnProSerTyrProAsnHisLysSerValPheAsnSerAla	140
Db	427	GCCACTGTCTGGATCTCGAAGCCCGGACCTGGCGGACCACTCTGCGATCGTGAATATCTC	486

141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 487 CGCATCCCGATCGGGAATACCGTATTTTCAGACCGGAGCGCGCGCGCGCGCGCG 546
 161 AlaLeuLeuAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 547 GCGTTCATCGAGGATCTGCCGAGGTAAGGCGGCGGCGGCGGCGGCGGCGGCGG 606
 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 607 TCCCAACCCGACCGCGCCCAACCGACCGCGGTCAGTGGCTGCGCGGAGCG 666
 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaGly 220
 667 CTGGCCCGGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 726
 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisGlyLeuLeuVal 240
 727 CTCGAGATGATCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 786
 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 787 GCGGCTCTCTGCTCGAAGAACCTTCGCACTACCGGCGGCGGCGGCGGCGGCGGCGG 846
 261 Val-----AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAla 278
 847 ATCGCGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 897
 279 IleArgAlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeu 298
 898 AACCGCGCAGAACTACTCTCTTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 957
 299 SerAsnAspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIle 318
 958 GAGGACGAGACGCTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1017
 319 GlnArgMetArgGlnLeuPheValAsnThrLeuGln---GluLysGlyAlaAsnArgAsp 337
 1018 CTGACGCTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG --- 1074
 338 PheSerPheIleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGln 357
 1075 TTGCGGCTCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1134
 358 ValLeuArgLeuArgGluPheGlyValTyrAlaValAlaSerGlyArgValAsnVal 377
 1135 GTGGAGCGGCTCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1194
 378 AlaGlyMetThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 1195 GCGGGGCTGAACCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1251

RESULT 8
 US-10-979-821-1
 ; Sequence 1, Application No. US10979821
 ; Publication No. US2005024937A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ABRAHAM, TIMOTHY W.
 ; APPLICANT: CAMERON, DOUGLAS C.
 ; APPLICANT: HICKS, PAULA M.
 ; APPLICANT: MCFARLAN, SARA C.
 ; APPLICANT: MILLIS, JIM
 ; APPLICANT: ROSAZZA, JACK
 ; APPLICANT: ZHAO, LISHAN
 ; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
 ; PRODUCTION OF MONATIN AND ITS PRECURSORS
 ; FILE OF INVENTION: 023829-0390
 ; CURRENT APPLICATION NUMBER: US/10/979,821
 ; CURRENT FILING DATE: 2004-11-03
 ; PRIOR FILING DATE: 2003-04-23
 ; PRIOR APPLICATION NUMBER: 60/374,831

102-04-23
 ; NUMBER OF SEQ ID NOS: 77
 ; SOFTWARE: PatentIn Ver. 3.3
 ; SEQ ID NO 1
 ; LENGTH: 1170
 ; TYPE: DNA
 ; ORGANISM: Sinorhizobium meliloti
 ; US-10-979-821-1
 Alignment Scores:
 Pred. No.: 8 93e-79 Length: 1170
 Score: 815.00 Matches: 170
 Percent Similarity: 60.7% Conservative: 63
 Best Local Similarity: 44.3% Mismatches: 137
 Query Match: 39.9% Indels: 14
 DB: 8 Gaps: 4
 US-10-673-786A-2 (1-396) x US-10-979-821-1 (1-1170)
 QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProfileLeuGlyLeuAlaAspLeuPhe 20
 DB 1 ATGTTGACGCGCTCGCCCGCAAGCCGACGATCCCTTCTTCTGATCGCGCTGTTTC 60
 QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 DB 61 AGGAAGGATGAGCGCGCGGAAAGTGCATCTCGCGTAGGAGTCTATCGCGACGAGACC 120
 QY 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTrpLeuLeuGluAsnGlu 60
 DB 121 GGACACGCGCGATCTTCGCGCGCTCAAGCGCGGAAAGCGCTTCTCGAAACACAG 180
 QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 DB 181 GACACGAGCGCTATATCGCCCGGAGGAGGACCTCTTCTTCGATCGGCTCTGGAA 240
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuLeuAsnAspLysArgAlaArgThrAlaGlnThr 100
 DB 241 CTGCTC-----GGCGGCGACACGATCGAGCGGAGCCATGTTGCGGCGCTCCAGACG 291
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaAsnThrSerVal 120
 DB 292 CCGCGCGCTCGCGCGCTCGTTCGCGCGGAGACCTCATCGCCCGC---ATGCGCGCGC 348
 QY 121 LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 DB 349 CGAGCATCTGCTCGCGCTCGGAGCTGGCGGAGACACCGCGCGCGGATCTTCAAGCGCGC 408
 QY 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 DB 409 GGGCTCGATATCGCCACCTACGACCTTTCGACATTCGCTCGCAGTCGCTCATCTTCGAT 468
 QY 161 AlaLeuLeuAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 DB 469 AATCTGGTGAAGCGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 528
 QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 DB 529 TGCCCAACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 588
 QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 DB 589 GTGGCGGAGCGGCGGCTGCTCGCGCTCGTGCATCTCGCTATCAGGGGTTCGCGCGCGC 648
 QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuLeuVal 240
 DB 649 CTCGACGAGGATGTCGCGGCGCTCGGCGATCTTCTCGCGGTGGTCCCGGAGCGGCTC 708
 QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 DB 709 GCGGTTCTCTGCTCGAAGTCTCTCGGCTTATTCGCGAGCGGCGGCGGCGGCGGCTCT 768
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280

Db 769 CGGACCACTCGACTCGCTCGCGCGAGCAGGGTGGCTCAAACTCGCGGGCTCGCAGCG 828
Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValAlaAlaThrIleLeuSerAsn 300
Db 829 ACCAGCTATTCCATCCCGCGGATCACCGCGAGCGCTGTCGGACGATCTTCAGCAGC 888
Qy 301 AspAlaLeuArgAlaIleTrrpLugluGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 889 CCGGAACCTCAGCGCGACTGGACGCGAGGAGCTCGAGACGATCGCGCTCAGGATGACGGC 948
Qy 321 MetArgGlnLeuPheValAsnThrLeu-----GlnGluLysGlyValAsnArg 336
Db 949 CTCGCGCGTCCGCTTCGCGAGGACTCGCACCCGCTCGCAGAGCTCGCGCGAGTCCGCC 1008
Qy 337 AspPheSerPheIleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGlu 356
Db 1009 GAT-----CAGGAGGCGATGTTCTCCATGCTGCCGCTTCGGAAGCG 1050
Qy 357 GlnValLeuArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsn 376
Db 1051 GAGGTTATCGCGCTCAGGACCGAGCAGCGCATCTATATCGCGCATCGCGCGCATCAAC 1110
Qy 377 ValAlaGlyMet 380
Db 1111 ATCGCGGGCTG 1122

RESULT 9

US-11-114-922-1
; Sequence 1, Application US/11114922
; Publication No. US20050282260A1
; GENERAL INFORMATION:
; APPLICANT: MCFARLAN, SARA C.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; FILE REFERENCE: PRODUCTION OF MONATIN AND ITS PRECURSORS
; CURRENT APPLICATION NUMBER: US/11/114,922
; CURRENT FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 1
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Sinorhizobium meliloti
US-11-114-922-1

Alignment Scores:
Pred. No.: 8,938-79 Length: 1170
Score: 815.00 Matches: 170
Percent Similarity: 60.7% Conservative: 63
Best Local Similarity: 44.3% Mismatches: 137
Query Match: 39.9% Indels: 14
Gaps: 4

US-10-673-786A-2 (1-396) x US-11-114-922-1 (1-1170)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProfileLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTTGAGCCCTCGCCCGCCAGCCGACGATCCCTTCTTCCTGATCGGCTGTTTC 60
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 61 AGGAAGGATGAGCGCGCGGAAAGTCGATCTCGCGTAGGAGTCTATCGCGACGAGACC 120
Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 121 GGACGACGCGCATCTTCGGGCGCTCAAGGCGGGAAGCGGCTTCTCGAAACACAG 180
Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80

Db 181 GACAGCAAGCCCTATATCGGCCCGAAGGGAGCACTCGTCTTTCTCGATCGGCTCTGGAA 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaAthrAlaGlnThr 100
Db 241 CTGCTC-----GGCGGGACACGATCGAGCGGAGCCATGTTGGGGGCTCCAGACG 291
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 292 CCGGCGGCTCGCGCGCTCCGTTTGGCGCGACCTCATCGCCCGC--ATGGCGCGC 348
Qy 121 LysArgValTrrpValSerAsnProSerTrrpProAsnHisLysSerValPheAsnSerAla 140
Db 349 CGAGGCATCTGGCTCGCGCTCGCGAGTGGCCGCAACCAACCGCGCATCTTCAAGCGCGCC 408
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 409 GGGCTCATATCGCCACTAGCACTTCTTCGACATTCGCTCGCAGTCGGTTCATCTTCGAT 468
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 469 AATCTGTGAGCGCGCTGGAAGCGCGCATCGCGCATCGGTGCTGCTCATGCAAGC 528
Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrrpGlnThrLeuAlaGlnLeu 200
Db 529 TGCCACAACCCGACCGCGCGCTCCTCGAGCGAAGCAACAATGGATGGAGTCCGCGCTG 588
Qy 201 SerValGluLysGlyTrrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 589 GTGCCCGAGCGCGCTGCTGCGCTCGTCTGATCTCGCTATCAGGGGTTTCGGCGCGCGC 648
Qy 221 LeuGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGlyLeuIleVal 240
Db 649 CTGACCCAGGATGTCGCGCGCTCCGCGCATCTTCTCGCGCTGGTCCCGGAAGCGCTCGTC 708
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 709 GCGGTTTCTGCTCGAAGTCTTTCGGCTTTATCGGAGCGCGCGCGCGCATCTTCGCG 768
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 769 CGGACCACTCGACTGCTCGCGGACAGGCTGCGCTCAAACTCGCGGGCTCGCAGCGC 828
Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 829 ACCAGCTATTCCATGCGCGCGATCACGGCGAGCGCTCGTGGCGACGATCCTTGACGAC 888
Qy 301 AspAlaLeuArgAlaIleTrrpGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320
Db 889 CCGGAACCTCAGCGCGACTCGACCGAGAGGCTCGAGACGATCGCGCTCAGGATCGGCGC 948
Qy 321 MetArgGlnLeuPheValAsnThrLeu-----GlnGluLysGlyValAsnArg 336
Db 949 CTCGCGCGCTCGCTTCGCGAGGAGCTCGCACCCGCTGGCAGAGCTTCGGCGAGTCCGCC 1008
Qy 337 AspPheSerPheIleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGlu 356
Db 1009 GAT-----CAGGAGGCGATGTTCTCCATGCTCGCGCTTTCGGAAGCG 1050
Qy 357 GlnValLeuArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsn 376
Db 1051 GAGGTTATCGCGCTCAGGACCGGACGCGCATCTATATCGCGCATCGCGCGCATCAAC 1110
Qy 377 ValAlaGlyMet 380
Db 1111 ATCGCGGGCTG 1122

RESULT 10

US-10-979-821-3
; Sequence 3, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.


```

US-10-673-786A-2 (1-396) x US-11-136-527-2278 (1-1260)
Qy 1 MetPheGluAenIleThrAlaAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Db 76 ATGTCAGCGCCCTGAAGCCGAGCCCGGAGCAAGATCCTGCAACTGATCCAGATGTC 135
Qy 21 ArgAlaAaspGluArgProGlyLysIleAenLeuGlyIleGlyWalTyLysAaspGluThr 40
Db 136 CGCAGGATGCGCGCGGACAAAGATCGATCTGGCGGTGGCGCTTACAAGACCCGAC 195
Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 196 GGGCTACCCCGCTGTCGGGCGCTGAAGCGCGCGAGAGCGGCTCTGGAGGTGCG 255
Qy 61 ThrThrLysAenTyrLeuGlyLysAaspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 256 ACCACCAAGACCTACACCGCCTCCGCGAGCGCGCTACATGTCGCGATGCGGAAG 315
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAaspLysArgAlaArgThrAlaGlnThr 100
Db 316 CTGATCCTC-----GCGGCGCGGTCCGCGCGACCGCGGTGCGCTCGGTCGCCAC 366
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAaAaspPheLeuAlaLysAenThrSerVal 120
Db 367 CCGCGCGGACCGCGCGGTGCGTTCAGCGCTCGAGTCCGATCGCGTCCGCGCGAG 426
Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 427 GCCACCGCTCGATCTCGAACCGACCTGCGGACCATCTGTCGATCGTGAATATCTC 486
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTrpAapAlaGluAenHisThrLeuAaspPheAasp 160
Db 487 GGCATCCCGGACCGGGAATACCGTATTTCACGCGCGAGACCGCGCGCTCGATGCCG 546
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCys 180
Db 547 GGCATGATGGAGGATCTGCCGAGTGAAGCGCGCGACGCTGCTGCTGACGCGGTGC 606
Qy 181 CysHisAenProThrGlyLysAaspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 607 TGCCACACCGACCGCGCCCAACCGACCGGTGTCAGTGGTGGCCATCTCGGAGAGC 666
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAaspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 667 CTGGCCCGGACAGCGCGCGTGGCTGATCGACCTCGCTATACAGGCTTCGGCGACGG 726
Qy 221 LeuGluGluAaspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 727 CTCGAGATGGATGCGCGCGGACGCGGTCTTGGCCACCGACTCCCGGAGGTGCTGATC 786
Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyAlaCysThrLeu 260
Db 787 GCGGCTCTCTCGAAGAACTTCGGCATCTACCGGAGCGCACCGGGATCTGATCGCC 846
Qy 261 Val-----AlaAlaAaspSerGluThrValAaspArgAlaPheSerGlnMetLysAla 278
Db 847 ATCGCGGAGCGCGCGCGGACGCGTGCAG-----GCCAACCTCAACTTCCTG 897
Qy 279 IleArgAlaAenTyrSerAenProAlaHisGlyAlaSerValValAlaThrIleLeu 298
Db 898 AACCGCGAGAACTACTCTTCGCGCGGACCATGCGCGCGCTCGTGACCATGATCTC 957
Qy 299 SerAenAaspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAaspMetArgGlnArgIle 318
Db 958 GAGGACGAGACGCTGAGCGCGACTGGAGCGGGAACCTCGAGAGGTGCGCTCAACATG 1017
Qy 319 GlnArgMetArgGlnLeuPheValAenThrLeuGln---GluLysGlyAlaAenArgAasp 337
Db 1018 CTGACACTGCGCGCGCAGCTTGGCGTTCGAGCGCGAGACCGGCTCGAACCGC--- 1074
Qy 338 PheSerPheIleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGln 357
Db 1075 TTCGGCTTCGTGGCGGACATCGCGGCATGTTCTCGCGCCTCGGGATCACCGCCCGAG 1134

```

```

Qy 358 ValLeuArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenVal 377
Db 1135 GTGAGCGGCTGCGGACCGGACGCGGCTTACATGTTGGCGCATTCGCGCTGAACATC 1194
Qy 378 AlaGlyMetThrProAaspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1195 GCGGGGCTGAACCGGACGCGTCCGCGTCTGCGCGCGCGGTGGCCAAAGTGTCTG 1251

```

RESULT 12

```

US-11-136-527-2278
; Sequence 2278, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; PRIOR FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2278
; LENGTH: 2343
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2278

```

Alignment Scores:

Pred. No.:	1.08e-74	2343
Score:	781.00	162
Percent Similarity:	57.3%	62
Best Local Similarity:	41.4%	165
Query Match:	38.2%	2
DB:	12	2

US-10-673-786A-2 (1-396) x US-11-136-527-2278 (1-2343)

```

Qy 4 AsnIleThrAlaAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPheArgAlaAasp 23
Db 192 CATGTTGAAATGGGACCTCCAGATCCCATCTCGGAGTGACCGAGGCTTCAGAGAGAT 251
Qy 24 GluArgProGlyLysIleAenLeuGlyIleGlyValTyrLysAaspGluThrGlyLysThr 43
Db 252 ACCAACAGCAGAAAGATGAACCTGGGAGTTGGTGCCTACCGGACGATAACGGAAGCCT 311
Qy 44 ProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAenGluThrThrLys 63
Db 312 TACGTGCTCCCAAGTGTTCGGAAGGACAGAGCCAGATTCCTGGGAAAAAATTTGGWCAA 371
Qy 64 AsnTyrLeuGlyLysAaspGlyIleProGluPheGlyArgCysThrGlnGluLeuPhe 83
Db 372 GATACCTTACCATCGGGAGCTGGCTGATTTTGTAAAGGCTCTCGAGAACTGGCCCTG 431
Qy 84 GlyLysGlySerAlaLeuIleAenAaspLysArgAlaAargThrAlaGlnThrProGly 103
Db 432 GCGGAGAACGCGAAGTGTGAAAAAGCGCGGTTTGTAACTGTGACAGACCATTTCCGG 491
Qy 104 ThrGlyAlaLeuArgValAlaAlaAaspPheLeuAlaLysAenThrSerValLysArg--- 122
Db 492 ACTGAGCCTTGAGGGTTCGAGCGCAGCTTTCTGCAAAAGATTTTTTAAGTTCAGCGCAG 551
Qy 123 ValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAlaGlyLeu 142
Db 552 GTCTTTCTGCCAACCATCTCTGGGAAACACACAGCCCATCTTCAGGGATGCGCGCATG 611
Qy 143 GluValArgGluTyrAlaTyrTrpAaspAlaGluAenHisThrLeuAaspPheAaspAla 162
Db 612 CAGCTGCAAGGTTATCGATACTATGACCCCAAGACTTGGCGCTTTGAGCTTCTCTGGAGCC 671
Qy 163 IleAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCysHis 182

```

```

Db      672  TTAAGAGACATATCAAAATCCAGAGAGAGTGTCTCTCTTGTGACGCGCTCAC 731
Qy      183  AenProThrGlyLeuAaspProThrLeuGluGlnThrGlnThrLeuAlaGlnLeuSerVal 202
Db      732  AACCCACGGCGTGGAGCCGCTCCAGAGAGTGGAGGAATGGCGCGTGGTGAAG 791
Qy      203  GluLysGlyTrpLeuProLeuPheAaspPheAlaTyrglnGlnPheAlaArgGly---Leu 221
Db      792  AAAAGAATCTCTCGCATCTTTGACATGGCTACCAAGGCTTTGCCAGCGCGCATGT 851
Qy      222  GluGluAaspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGlnLeuLeuVal 241
Db      852  GATAAGGAGCGCTGGCGCGCTTCATCGAGCGAGGCAATCAATGTCTGCGCTGCG 911
Qy      242  SerSerTyrsLysAenPheGlyLeuTyraenGluAatGValGlyValaCysThrLeuVal 261
Db      912  CAATCTATGCCAAGAACATGGCCCTGTACCGTGAGCGGTGGAGCGCTTCACTGTGTS 971
Qy      262  AlaAlaAaspSerGluThrValAaspArgAlaPheSerGlnMetLysAlaAlaAla 281
Db      972  TCAAGATGCAAGAGAACCAAGGGTGGAGTCACAGCTCAAGATCTGTATCGCCCC 1031
Qy      282  AenTyrsAenProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAenAep 301
Db      1032  TTGTATTCAACCGCGCTCTCAATGAGCGCGATCGCGCAACCATYCTGAYKCTSSA 1091
Qy      302  AlaLeuArgAlaIleTrpGlnGlnLeuThrAaspMetArgGlnArgIleGlnArgMet 321
Db      1092  GACKXSCGAAGCAATGGTGTGAGGAGGTGAAGCGATGGCTGACGCGATCATCAGCATG 1151
Qy      322  ArgGlnLeuPheValaenThrLeuGlnGluLysGlyAlaAenAaspPheSerPheIle 341
Db      1152  AGACCCAGTGTCTTCAACCTGAAGAAAGAGGCTGTGCCACAACTGGCGAGCACATC 1211
Qy      342  IleLysGlnAasnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeu 361
Db      1212  ACCGACCGATCGGATGTTCTGCTTACCGCGCTAAAGCGCTGAGCGGTGGAGCGGTG 1271
Qy      362  ArgGluGluPheGlyValTyraValAlaLaserGlyArgValaenValaGlyMetThr 381
Db      1272  ACCAAGAGKCTYAGKCTACATGACAAAGAGTGGTGGATCTCTGKGGCGGGGTGACC 1331
Qy      382  ProAaspAenMetAlaProLeuCysGluAlaIle 392
Db      1332  TCTGCAATGTGGGTAMCTGGCCCAAGCCATT 1364

```

```

RESULT 13
; Sequence 82350, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHITHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHKART, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82350
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-82350

Alignment Scores:
Pred. No.: 3,43e-69 Length: 1251
Score: 727.50 Matches: 153
Percent Similarity: 57.5% Conservative: 74

```

```

Best Local Similarity: 38.7% Mismatches: 157
Query Match: 35.6% Indels: 11
DB: 7 Gaps: 6
US-10-673-786A-2 (1-396) x US-10-932-182A-82350 (1-1251)
Qy      1  MetPheGluAenIleThrAlaAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Db      55  CTGTTCAATAACATCGAATGCTGCCCTGATGCCCTTTTGGTATTAAAGCAAGGTAC 114
Qy      21  ArgAlaAaspGluAatGProGlyLysIleAenLeuGlyIleGlyValTyrsAepGluThr 40
Db      115  GGGCAAGATCAACGGTCTACCAAGGTGACCTGGGTATCGGGCCCTACAGACGACCAAC 174
Qy      41  GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrsLeuLeuGluAenGlu 60
Db      175  GGTAAACCATGGGTCTTCCCAAGTGTAAAGCCGCGGAAAG---CTAATTCATAACGAC 231
Qy      61  ThrThrLysAen-----TyrLeuGlyIleAaspGlyIleProGluPheGlyArgCysThr 78
Db      232  AGCTCTCAACACCATGAATACCTCGGTATTACCGGTCTGCCAAGTTTGACATCTAAACGCC 291
Qy      79  GlnGluLeuLeuPheGlyLysGlySerAlaLeuIleAenAaspLysArgAlaArgThrAla 98
Db      292  GCCAAGATCATCTTCGGTACCGAATCCGATCGCTTGCAGGAAGACAGAGTAATCTCAGTA 351
Qy      99  GlnThrProGlyGlyThrGlyValAlaLeuArgValAlaAaspPheLeuAlaLysAenThr 118
Db      352  CAATCAGCTGTGTGACGGGTCTCTCATATATTGGGAAGTTTTTTCAAATTTCTTC 411
Qy      119  SerValLysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAen 138
Db      412  CCAGATAAATCGTCTATTGCTAAGCTACTTGGGCCAACCAACATGCGCCATTTTGTAG 471
Qy      139  SerAlaGlyLeuGluValArgGluTyraTyrsAaspAlaGluAenHisThrLeuAap 158
Db      472  AATCAAGCTTGAAGAACCGCGACTTACCTTTACTCGGCCCAACGAACTAAAGCTTTGGAC 531
Qy      159  PheAaspAlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHis 178
Db      532  CTAAACGGCTTCTAAATGCTATTCAAAGAGCTCCAGAGGCTCCATTTTCTGTCGAC 591
Qy      179  GlyCysCysHisAenProThrGlyIleAaspProThrLeuGluGlnTrpGlnThrLeuAla 198
Db      592  TCTTGGCGCCATAACCCAACTGGTCTGGACCTACTAGTGAACAATGGGTTCAAATCGTT 651
Qy      199  GlnLeuSerValGluLysGlyTrpLeuProLeuPheAaspPheAlaTyrglnGlyPheAla 218
Db      652  GATGCTATCGCTCAAAAAACCAATCGCTTATTGTACACGCGCTTACCAAGGTTTGGC 711
Qy      219  ArgGly---LeuGluGluAaspAlaGluGlyLeuArgAlaPheAla-----AlaMet 234
Db      712  ACTGGAGATTGGACAAGGATGCTTGTGCGCTCTAGGTGTGGAGAGCTTTCGACG 771
Qy      235  HisLysGluLeuIleValAlaSerSerTyrsLysAenPheGlyLeuTyrsAenGluArg 254
Db      772  GTCCTCCCGCTTGTGTCTGTCAGTCTTTCGCAAGAACCGCGGTATGTACGCTGAGCGT 831
Qy      255  ValGlyAlaCysThrLeuVal-----AlaAlaAaspSerGluThrValAaspArgAla 271
Db      832  GTAGGTTGTTTCCATCTAGCACTTACAAAAACAAGCTCAAAACCAAACTATAAAGCGCTGCT 891
Qy      272  Phe---SerGlnMetLysAlaAlaIleArgAlaAenTyrsAenProProAlaHisGly 290
Db      892  GTTACATCTCAATTGGCCAAAATCATCTGTAGTGAAGTGTCCAAACCCACCGCGCTACGCG 951
Qy      291  AlaSerValValAlaThrIleLeuSerAenAaspAlaLeuArgAlaIleTrpGluGlnGlu 310
Db      952  GCTAGATTGCTGCTAAACTGTTGAAACGCCAGATTAACGGAACAGTGGGCAAGGAT 1011
Qy      311  LeuThrAaspMetArgGlnArgIleGlnArgMetArgGlnLeuPheValaenThrLeuGln 330
Db      1012  ATGGTTACCATGTCTCTCCAGAATTACGAAAATGAGCGCAGCATTAAGAGACCATTTAGTC 1071

```


Alignment Scores:

Pred. No.: 1,77e-68 Length: 1389
Score: 721.50 Matches: 155
Percent Similarity: 55.8% Conservative: 71
Best Local Similarity: 38.3% Mismatches: 170
Query Match: 35.3% Indels: 9
DB: 7 Gaps: 5

US-10-673-786a-2 (1-396) x US-10-932-182A-4671 (1-1389)

```
Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
   |||||
Db 145 CTATTCAACAACATCGTCTCCCGCTGACGCCCTTTTGGGATCAAGCAAGATAC 204
   |||||
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrlAspGluThr 40
   |||||
Db 205 GGTCAAGACCAACGACCAACCAAGCTCGACCTGGGTATCGGAGCCTACAGACAAATAAC 264
   |||||
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrlLeuLeuGluAsnGlu 60
   |||||
Db 265 GGTAAAGCCATGGGTCTCTGCAAGTGTCAAAGCCGCGGAAAGCTTATCCATAGCGACGAA 324
   |||||
Qy 61 Thr----ThrLysAsnTyrlLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGln 79
   |||||
Db 325 TCCTACAATACAGAGTACCTCGGATCACTGGGCTGCCCTCTTTGACGTCCAACGCCGCC 384
   |||||
Qy 80 GluLeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGln 99
   |||||
Db 385 AAGATCATCTTCGGTACGCAATCCCGCCCTTTCAGAGGAGGACAGAGTAATTTTCGGTGCAG 444
   |||||
Qy 100 ThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSer 119
   |||||
Db 445 TCGTGTCCGGTACTGGAGCCCTTCATATAGCCGCCAAGTCTTCTCCAAGTTCTATCCA 504
   |||||
Qy 120 ValLysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSer 139
   |||||
Db 505 GACAACTGGTCTACCTGTCTCAAGCCACCTCGGCAAAACCAACACGCCCATCTTTGAAAAC 564
   |||||
Qy 140 AlaGlyLeuGluValArgGluTyrlAlaTyrlAspAlaGluAsnHisThrLeuAspPhe 159
   |||||
Db 565 CAAGGTTTGAACAAACCAACCAACCACTACCCCTACCTGGGCCAATGAGCAAGTCCTTGGATT 624
   |||||
Qy 160 AspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGly 179
   |||||
Db 625 GACGGCTTCTCGCGGCATCAAAAGCCGCCAGAGGGTTCGCTCTTTGTTCTTCACTCT 684
   |||||
Qy 180 CysCysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGln 199
   |||||
Db 685 TGTGCTCAACACCTACTGTCTGGACCACTAACACCAATGGGCCCAAGATTGTCGAC 744
   |||||
Qy 200 LeuSerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrlGlnGlyPheAlaArg 219
   |||||
Db 745 GCCATTGGCCGCAAGAACCAACCGTCTCCCTATTTCGACACCGCGTACAGGGTTTCGCTACA 804
   |||||
Qy 220 Gly----LeuGluGluAspAlaGluGlyLeuArgAlaPheAla-----AlaMethis 235
   |||||
Db 805 GGTGACCTTGGCAAGAACGCTTACCCGCTGGCTCTGGGTGGGAGAAACTATCCACAGTC 864
   |||||
Qy 236 LysGluLeuIleValAlaSerSerTyrlLysAsnPheGlyLeuTyrlAsnGluArgVal 255
   |||||
Db 865 TCTCCCGTCTTGTCTGCCAGTCTGTTGCCAAGAACGCTGGTATGTACGGTGAGCGTGTG 924
   |||||
Qy 256 GlyAlaCysThrLeuVal-----AlaAlaAspSerGluThrValAspArgAlaPhe 272
   |||||
Db 925 GGTGTTTCCATCTACGCGCTCAAAAGCAAGCCAAACCAAGACCATCAAGCCCGCTGTC 984
   |||||
Qy 273 ---SerGlnMetLysAlaAlaIleArgAlaAsnTyrlSerAsnProAlaHisGlyAla 291
   |||||
Db 985 ACCTCCCAATTGGCCCAAGATCATAGTAGTGAAGTGTCCATTCACCCCGCTTACGGTGCC 1044
   |||||
Qy 292 SerValValAlaThrIleLeuSerAsnAspAlaLeuArgAlaIleTrpGluGlnGluLeu 311
   |||||
```

Search completed: March 15, 2006, 16:14:48

Job time : 466 secs

```
Db 1045 AAGATCGTCTAGCTGTTGGAACACCCGGAATTGACCGCAATGCGACAAATGGCACAGGATATA 1104
Qy 312 ThrAspMetArgGlnArgIleGlnArgMetArgGlnLeuPheValAsnThrThrLeuGlnGlu 331
   |||||
Db 1105 GTCACCATGTCTCCAGATCACCACAAATGAGACATGCTCTAAGAGAGACCACCTTGGTTAAG 1164
   |||||
Qy 332 LysGlyAlaAsnArgAspPheSerPheIleIleLysGlnAsnGlyMetPheSerPheSer 351
   |||||
Db 1165 CTGGCCACCCCGGGCAATTGGGACCATATAGTCCACCAAGTGTGGGATGTTCTCTTCACC 1224
   |||||
Qy 352 GlyLeuThrLysGluGlnValLeuArgLeuArgGluGluPheGlyValTyrlAlaValAla 371
   |||||
Db 1225 GGGTTGACTCTCTATATGTCAGAGACTAGAAGAAACCCACGCCGTTTACTTGTGTCGCC 1284
   |||||
Qy 372 SerGlyArgValAsnValAlaGlyMetThrProAspAsnMetAlaProLeuCysGluAla 391
   |||||
Db 1285 TCGGGCAGAGCTCCATTGGCGGGTTGAACCAAGGTAACTGGAATACGTAGCCCAAGGCC 1344
   |||||
Qy 392 IleValAlaValLeu 396
   |||||
Db 1345 ATTGACGAAGTGGTG 1359
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 15:34:18 ; Search time 6793 Seconds
(without alignments)
3223.239 Million cell updates/sec

Title: US-10-673-786A-2
Perfect score: 2045
Sequence: 1 MFENITAPADPILGLADLF.....VAGMTPDNNAIPLCAIVAVL 396

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 79147668 seqs, 27645789525 residues

Total number of hits satisfying chosen parameters: 158295336

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame+ p2n.model -runat 14032006 135615 22715/app_query.fasta 1
-Q=/abss/ABSSWEB_spool/US10673786 -SUFFIX=rlnmp -MINMATCH=0.1 -LOOPCL=0
-DB=Pending Patents NA Main -QFMT=fastap -SUFFIX=rlnmp -TRANS=human40.cdi
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -ALIGN=15
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss04 -USER=US10673786 @CGN 1 1 7336 @runat 14032006 135615 22715
-NCPU=3 -ICPU=3 -NO MAP -NEG SCORES=0 -WAIT -DSBLOK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA Main:
1: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
2: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
3: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
4: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
5: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
6: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
7: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
8: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
9: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
10: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
11: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
12: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
13: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
14: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
15: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
16: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
17: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
18: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
19: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
20: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
21: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
22: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
23: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
24: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
25: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
26: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
27: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
28: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:

29: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:
30: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:
33: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:
34: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:
35: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:
36: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:
37: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:
38: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:
39: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:
40: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:
41: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:
42: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:
43: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:
44: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:
45: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:
46: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:
47: /cgn2_6/ptodata/1/pna/US103C_COMB.seq:
48: /cgn2_6/ptodata/1/pna/US103D_COMB.seq:
49: /cgn2_6/ptodata/1/pna/US103E_COMB.seq:
50: /cgn2_6/ptodata/1/pna/US103F_COMB.seq:
51: /cgn2_6/ptodata/1/pna/US104A_COMB.seq:
52: /cgn2_6/ptodata/1/pna/US105A_COMB.seq:
53: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:
54: /cgn2_6/ptodata/1/pna/US106B_COMB.seq:
55: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:
56: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:
57: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:
58: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:
59: /cgn2_6/ptodata/1/pna/US107E_COMB.seq:
60: /cgn2_6/ptodata/1/pna/US107F_COMB.seq:
61: /cgn2_6/ptodata/1/pna/US107G_COMB.seq:
62: /cgn2_6/ptodata/1/pna/US108A_COMB.seq:
63: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:
64: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:
65: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:
66: /cgn2_6/ptodata/1/pna/US110A_COMB.seq:
67: /cgn2_6/ptodata/1/pna/US110B_COMB.seq:
68: /cgn2_6/ptodata/1/pna/US110C_COMB.seq:
69: /cgn2_6/ptodata/1/pna/US110D_COMB.seq:
70: /cgn2_6/ptodata/1/pna/US111A_COMB.seq:
71: /cgn2_6/ptodata/1/pna/US111B_COMB.seq:
72: /cgn2_6/ptodata/1/pna/US112A_COMB.seq:
73: /cgn2_6/ptodata/1/pna/US117A_COMB.seq:
74: /cgn2_6/ptodata/1/pna/US600A_COMB.seq:
75: /cgn2_6/ptodata/1/pna/US601A_COMB.seq:
76: /cgn2_6/ptodata/1/pna/US602A_COMB.seq:
77: /cgn2_6/ptodata/1/pna/US602B_COMB.seq:
78: /cgn2_6/ptodata/1/pna/US603A_COMB.seq:
79: /cgn2_6/ptodata/1/pna/US604A_COMB.seq:
80: /cgn2_6/ptodata/1/pna/US604B_COMB.seq:
81: /cgn2_6/ptodata/1/pna/US605A_COMB.seq:
82: /cgn2_6/ptodata/1/pna/US606A_COMB.seq:
83: /cgn2_6/ptodata/1/pna/US607A_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2045	100.0	1191	50	US-10-369-493-24472
2	2045	100.0	1191	54	US-10-673-786A-1
3	2045	100.0	1191	78	US-60-360-039-24472
4	2045	100.0	1191	82	US-60-638-099-7
5	2045	100.0	14759	33	US-09-801-563-87
6	2045	100.0	14759	62	US-10-893-671-87
7	2045	100.0	4639675	83	US-60-709-960-1
8	2039	99.7	1576	82	US-60-643-717-112
					Sequence 24472, A
					Sequence 1, Appli
					Sequence 24472, A
					Sequence 7, Appli
					Sequence 87, Appli
					Sequence 1, Appli
					Sequence 112, App


```

9      2011      98.3      1191      70      US-11-114-922-75      Sequence 75, Appl
10     1931      94.4      1209      51      US-10-446-203-4902      Sequence 4902, Ap
11     1930      94.4      1209      22      US-09-252-691-1606      Sequence 1606, Ap
12     1930      94.4      1209      22      US-09-252-691C-1606      Sequence 1606, Ap
13     1930      94.4      1209      51      US-10-417-886-1606      Sequence 1606, Ap
14     1930      94.4      1218      22      US-09-252-691-1616      Sequence 1616, Ap
15     1930      94.4      1218      22      US-09-252-691C-1616      Sequence 1616, Ap
16     1930      94.4      1218      51      US-10-417-886-1616      Sequence 1616, Ap
17     1654      80.9      1188      50      US-10-369-493-44812      Sequence 44812, A
18     1654      80.9      1188      78      US-10-360-039-44812      Sequence 44812, A
19     1654      80.9      1191      71      US-11-150-804-4967      Sequence 4967, Ap
20     1654      80.9      1191      81      US-60-579-902-4967      Sequence 4967, Ap
21     1654      80.9      1191      81      US-60-579-902A-4967      Sequence 4967, Ap
22     1654      80.9      524707      71      US-11-150-804-14985      Sequence 14985, A
23     1654      80.9      524707      81      US-60-579-902-14985      Sequence 14985, A
24     1654      80.9      524707      81      US-60-579-902A-14985      Sequence 14985, A
25     1650      80.7      1191      72      US-11-216-782-3916      Sequence 3916, Ap
26     1650      80.7      1191      82      US-60-606-098-3916      Sequence 3916, Ap
27     1650      80.7      171354      72      US-11-216-782-16894      Sequence 16894, A
28     1650      80.7      171354      82      US-60-606-098-16894      Sequence 16894, A
29     1343      65.7      1191      82      US-60-691-214-2033      Sequence 2033, Ap
30     1343      65.7      7977      1      PCT-US03-17092-536      Sequence 536, App
31     1343      65.7      7977      2      PCT-US04-07001-536      Sequence 536, App
32     1343      65.7      7977      61      US-10-795-159-536      Sequence 536, App
33     1343      65.7      7977      79      US-60-453-134-536      Sequence 536, App
34     1343      65.7      7977      82      US-60-691-214-536      Sequence 536, App
35     1343      65.7      908766      2      PCT-US04-07001-685      Sequence 685, App
36     1343      65.7      908766      61      US-10-795-159-685      Sequence 685, App
37     1343      65.7      908766      82      US-60-691-214-685      Sequence 685, App
38     1343      65.7      1913428      82      US-60-691-214-772      Sequence 772, App
39     1325      64.8      1830121      14      US-08-426-787-1      Sequence 1, Appl
40     1325      64.8      1830121      49      US-10-329-670-1      Sequence 1, Appl
41     1325      64.8      1830121      64      US-10-981-687-1      Sequence 1, Appl
42     1318      64.4      966      50      US-10-369-493-23918      Sequence 23918, A
43     1318      64.4      966      78      US-60-360-039-23918      Sequence 23918, A
44     1318      64.4      4404      34      US-09-897-516-374      Sequence 374, App
45     1318      64.4      4404      34      US-09-897-516-1179      Sequence 1179, Ap

```

ALIGNMENTS

```

RESULT 1
US-10-369-493-24472
; SEQUENCE 24472, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24472
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-369-493-24472

```

```

Alignment Scores:
Pred. No.: 1.4e-314      Length: 1191
Score: 2045.00      Matches: 396
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match: 100.0%      Indels: 0
DB: 50      Gaps: 0

```

US-10-673-786A-2 (1-396) x US-10-369-493-24472 (1-1191)

```

Qy      1      MetPheGluAenlleThrAlaLaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Dy      1      ATGTTTGAACAATTATCCGCGCTCTCTGCGCAGCCGATTTCTGGCGCTGGCGATCTGTTT 60
Qy      21      ArgAlaAaspGluArgProGlyLysIleAsnLeuGlyLeuGlyValTyrLysAaspGluThr 40
Dy      61      CGTCCCGATGAAGCTCCGCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 120
Qy      41      GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAasnGlu 60
Dy      121      GGCNAACCCCGGTACTGACCGCTGAAAAGGCTGAAACAGTATCTGCTCGAAAATGAA 180
Qy      61      ThrThrLysAenTyrLeuGlyLysIleAaspGlyLeuProGluPheGlyArgCysThrGlnGlu 80
Dy      181      ACCACCAAAATTAACCTCGGCTATGACGGCATCTCCGAAATTTGGTCTGCTGACCTCAGGAA 240
Qy      81      LeuLeuPheGlyLysGlySerAlaLeuIleAsnAaspLysAlaAargThrAlaGlnThr 100
Dy      241      CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCACGGCAGACT 300
Qy      101      ProGlyGlyThrGlyAlaLeuArgValAlaAlaAaspPheLeuAlaLysAasnThrSerVal 120
Dy      301      CCGGGGGGCACTGGCGCACTACGGTGGCTGCCGATTTCTGGCAAAAATACCCAGCGTT 360
Qy      121      LysArgValTrpValSerAasnProSerTrpProAenHisLysSerValPheAasnSerAla 140
Dy      361      AAGCGTGTGGGTGAGCAACCCCAAGCTGCGCAACCATTAAGAGCGCTTTTAACTCTGCA 420
Qy      141      GlyLeuGlnValArgGluTyrAlaTyrTyrAspAlaGluAasnHisThrLeuAaspPheAasp 160
Dy      421      GGTCTGAAGTTCTGTAATACGCTATTATGATCGGAAATACACCTCTTGACTTCGAT 480
Qy      161      AlaLeuIleAasnSerLeuAasnGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCys 180
Dy      481      GCATGATTAAACGCTGAATGAAGCTCAGCTGGCGCAGCTAGTGTGCTGCTCCATGGCTGC 540
Qy      181      CysHisAasnProThrGlyLysIleAaspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Dy      541      TGCCATAACCAACCCGCTATCGACCTACGCTGGAACAATGGCAACACATGGCAACATC 600
Qy      201      SerValGlyLysGlyTrpLeuProLeuPheAaspPheAlaTyrGlnGlyPheAlaArgGly 220
Dy      601      TCCGTTTGAGAAAGCTGGTTACCGCTGTTGACTTCGCTTACCGAGGGTTTGGCCCGTGT 660
Qy      221      LeuGluGluAaspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Dy      661      CTGGAAGAAGATGCTGAAGGACTGCGGCGCTTTCGCGCTATGCATAAAGAGCTCATTTGT 720
Qy      241      AlaSerTyrSerLysAasnPheGlyLeuTyrAasnGluArgValGlyAlaCysThrLeu 260
Dy      721      GCCAGTTTCTACTCTAATAAACTTTGGCTGTGTACAAACGAGCGCTGTGGCGCTTGACTCTG 780
Qy      261      ValAlaAlaAaspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Dy      781      GTTGTCTCCGACAGTGAACCCGTTGATCGCGCATTCAGCCAAATGAAGCGCGGCTATTCGC 840
Qy      281      AlaAenTyrSerAasnProAlaHisGlyValaSerValValAlaThrIleLeuSerAasn 300
Dy      841      GCTAACTACTCTAACCCACGACGACGCGGCTTCTGTTGTTGCGCACCATCTCTGAGCAAC 900
Qy      301      AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAaspMetArgGlnArgIleGlnArg 320
Dy      901      GATCGTTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGCTATTTCAGCGT 960
Qy      321      MetArgGlnLeuPheValAasnThrLeuGlnLysGlyAlaAasnArgAaspPheSerPhe 340
Dy      961      ATGCGTCAGTGTGTTGTCGTAATACGCTCAGGAAAAGCGCGCAAAACCCGCGCTTTCAGCTTT 1020
Qy      341      IleIleLysGlnAasnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Dy      1021      ATCATCAACAGAACGCGCATGTTCTCTTCAGTGGCTGTGACAAAAGAACACAGTGTCTGCGT 1080

```

QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
DB 1081 CTGCGCAAGAGTTTGGCGTATATGCGTGTCTTCTGTCGCTAAATGTCGCGGATG 1140
QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaValAlaValLeu 396
DB 1141 ACACCAAGATACATGCTCGCTGTGCGAAGCATTTGTGGCAGTGCTG 1188

RESULT 2

US-10-673-786A-1

; Sequence 1, Application US/10673786A
; GENERAL INFORMATION:
; APPLICANT: AGHVERDIAN, VALERY ZAVENOVICH
; APPLICANT: SAVRASOVA, EKATERINA ALEKSEEVNA
; APPLICANT: KAPLAN, ALLA MARKOVNA
; APPLICANT: LOBANOV, ANDREY OLEGOVICH
; APPLICANT: KOZLOV, YURI IVANOVICH
; TITLE OF INVENTION: METHOD FOR PRODUCING L-THREONINE USING BACTERIA
; TITLE OF INVENTION: BELONGING TO THE GENUS ESCHERICHIA
; FILE REFERENCE: US-115
; CURRENT APPLICATION NUMBER: US/10/673,786A
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: PCT/JP03/02067
; PRIOR FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: 2002104983
; PRIOR FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1191)
US-10-673-786A-1

Alignment Scores:

Pred. No.: 1,4e-314 Length: 1191
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 54 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-673-786A-1 (1-1191)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
DB 1 ATGTTTGAGAACATTACCGCGCTCTCTGCGGACCGGATTCCTGGCGCTGGCGCATCTGTTT 60
QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
DB 61 CGTGGCGATGAACGTCGCGGCAAAATTAACCTCGGGATTTGGTGTCTATATAAGATGAGACG 120
QY 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
DB 121 GGCAAAACCCGGTACTACCGAGCGTGAAAGGCTGAACAGTATCTGCTCGAAATGA 180
QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
DB 181 ACCACCAAAATTAACCTCGCATTCACCGCATCCCTGAATTTGGTTCGCTGCACTCAGAA 240
QY 81 LeuLeuPheGlyLysGlySerAlaLeuLeuAsnAspLysArgAlaArgThrAlaGlnThr 100
DB 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCGACGCGCACAGCT 300
QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
DB 301 CCGGGGGGACCTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAATTAACCGCGTT 360
QY 121 LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140

DB 361 AAGCGTGTGTGGGTGAGCAACCCCAAGCTGGCCGAACCATTAAGAGCGCTCTTTAACTCTGCA 420
QY 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
DB 421 GGTCTGGAAAGTTTCGTAATACGCTTATATGATGCGGAAATACACACTCTTGACTTCGAT 480
QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
DB 481 GCACGTGATTAAACAGCTGAAATGAAGCTCAGGCTGCGAGCTAGTGTCTGTTCATGCTGCTC 540
QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
DB 541 TGCCATAACCCCAACCGGTATCGACCTACGCTGGAAACCAATGGCAACACTGGCACAACTC 600
QY 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
DB 601 TCCGTTGAGAAGGCTGGTTACCGCTGTTTGACTTCGCTTACAGGGTTTTCGCCGTGGT 660
QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
DB 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGGCGCTATGCATAAAGAGCTGATTGT 720
QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
DB 721 GCCAGTTCCTACTCTAAAAAATTTGCGCTGTACAAACGAGCGGTGTGGCGCTTGTACTCTG 780
QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
DB 781 GTTGCTCGGACAGTGAACCGTTGATCGGATTCAGCCAAATGAAGCGCGGATTCGC 840
QY 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValAlaThrIleLeuSerAsn 300
DB 841 GCTAACTACTGTAAACCCACGACACACGCGCTCTCTGTTGTCACCACTCTCTGAGCAAC 900
QY 301 AspAlaLeuArgAlaIleTyrGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320
DB 901 CATCGTTTACGTGCGATTTGGGAAACAAGAGCTGACTGATATCGCCAGCGGTATTCACGCT 960
QY 321 MetArgGlnLeuPheValAsnThrLeuGlnLysGlyValAsnArgAspPheSerPhe 340
DB 961 ATGCGTCAGTTGTTTCGTCAATACGCTGCAGGAAAGGCGCAACCCGACCTTCAGCTTT 1020
QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
DB 1021 ATCATCAACAGACGCGCATGTTCTCTCTCAGTGGCCCTGACAAAGAACAGCTGCTGCT 1080
QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
DB 1081 CTGCGCGAAGAGTTTGGCGTATATGCGGTTCGCTTCGCTGCGTAAATGTGCGCGGATG 1140
QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
DB 1141 ACACCAATAACATGGCTCCGCTGTGCGAAGCATTTGTGGCAGTGCTG 1188

RESULT 3

US-60-360-039-24472

; Sequence 24472, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24472
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli

US-60-360-039-24472

Alignment Scores:

Pred. No.: 1.4e-314 Length: 1191
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 78 Gaps: 0

US-10-673-786A-2 (1-396) x US-60-360-039-24472 (1-1191)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTTTGAGAACATTACCGCGCTCTGCCACCGGATTCGGCGCTGGCGGATCTGTTT 60
Qy 21 ArgAlaAspGluAArgProGlyVylleAenLeuGlyIleGlyValTyrLysAspGluThr 40
Db 61 CGTCCGATGAACGTCCTCCGGCAAAATTAACCTCGGGATTGGTGTCTATAAAGATGAGACG 120
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 121 GGCAAAACCCGGTACTGACCGAGGTGAAAAGGCTGACAGTATCTGCTCGAAATGAA 180
Qy 61 ThrThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 181 ACCACCAAAATTAACCTCGCATTTGACGGCATCCTGAAATTTGGTCGTCACCTCAGGAA 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaAArgThrAlaGlnThr 100
Db 241 CTGCTGTTTGGTAAGGTAGCGGCTGATCAATGACAAACGTCGTCACCGGCACAGACT 300
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 301 CCGGGGGCACTGGGGCACTAGCGGTGCTGCCGATTTCTTGGCAAAATTAACAGCGTT 360
Qy 121 LysArgValTrpValSerAsnProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 361 AAGCGTGTGGTGAGCAACCCCAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 420
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAenHisThrLeuAenPheAsp 160
Db 421 GGTCTGGAAGTTCTGTAATACGCTTATATGATCGGAAATCACTCTTTGACTTCGAT 480
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 481 GCATGTATACAGCTGAATGAAGCTCAGGCTGGCGACGTAGTGTCTTCATGGGTGC 540
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 541 TGCCATACCCCAACCGGTATACGCTACGCTGGAACAATGGCAACACACTGGCACAACTC 600
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 601 TCCGTTGAGAAGGCTGGTTTACCGCTGTTTGACTTCGCTTACCAGGGTTTTGCCCGGT 660
Qy 221 LeuGluGluAenAlaGluGlyLeuArgAlaPheAlaAenHisLysGluLeuIleVal 240
Db 661 CTGAAGAAGATGCTGAAGGACTGGCGCTTTCCGGCTATGCAATAAAGAGCTGATTGTT 720
Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluAArgValGlyAlaCysThrLeu 260
Db 721 GCCAGTCTCTACTCTAANAACCTTTGGCTGTACACGAGCGGTGTGGCGCTTGACTCTG 780
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 781 GTTGTGCCGACAGTGAACCGGTTGATCGCGCATTCAGCCAAATGAAAGCGGCGGATTCGC 840
Qy 281 AlaAenTyrSerAenProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
Db 841 GCTAACTACTCTAACCACACAGCACACGCGGCTTCTGTTGTGGCCACCATCTTAGGCAC 900
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320

Db 901 GATCGGTACGTGCGATTGGGAACAAGAGCTGACTGATATGCCCCAGCGTATTTCAGCGT 960
Qy 321 MetArgGlnLeuPheValAenThrLeuGlnGluLysGlyAlaAenArgAspPheSerPhe 340
Db 961 ATCGGTCAAGTTGTCGTCAATACGTCAGAGAAAAGCGCAAAACCGCAGACTTCAGGCTTT 1020
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1021 ATCATCAACAGAACGCGCATGTTCTCTCAGTGGCTGACAAAGAACAAGTCTGCGT 1080
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenValAlaGlyMet 380
Db 1081 CTGCGCAAGAGTTGGCGTATATGCGGTGCTTCTCGTCGCTAAATGTGGCGGATG 1140
Qy 381 ThrProAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1141 ACACAGATAACATGGCTCGCTGTGCGAAGCGATTCTGGCAGTGCTG 1188

RESULT 4

US-60-638-099-7

; Sequence 7, Application US/60638099

; GENERAL INFORMATION:

; APPLICANT: Edgerton, Michael D

; TITLE OF INVENTION: Transgenic Plants with Enhanced Agronomic Phenotypes

; FILE REFERENCE: 38-21(53720)

; CURRENT APPLICATION NUMBER: US/60/638,099

; CURRENT FILING DATE: 2004-12-21

; NUMBER OF SEQ ID NOS: 48056

; SEQ ID NO 7

; LENGTH: 1191

; TYPE: DNA

; ORGANISM: Escherichia coli

US-60-638-099-7

Alignment Scores:

Pred. No.: 1.4e-314 Length: 1191
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 82 Gaps: 0

US-10-673-786A-2 (1-396) x US-60-638-099-7 (1-1191)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTTTGAGAACATTACCGCGCTCTGCCACCGGATTCGGCGCTGGCGGATCTGTTT 60
Qy 21 ArgAlaAspGluAArgProGlyLysIleAenLeuGlyIleGlyValTyrLysAspGluThr 40
Db 61 CGTCCGATGAACGTCCTCCGGCAAAATTAACCTCGGGATTGGTGTCTATAAAGATGAGACG 120
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 121 GGCAAAACCCGGTACTGACCGAGGTGAAAAGGCTGACAGTATCTGCTCGAAATGAA 180
Qy 61 ThrThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 181 ACCACCAAAATTAACCTCGCATTTGACGGCATCCTCGCTGCTGCTGCTGCTGCTGCTG 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaAArgThrAlaGlnThr 100
Db 241 CTGCTGTTTGGTAAGGTAGCGGCTGATCAATGACAAACGTCGTCACCGGCACAGACT 300
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 301 CCGGGGGCACTGGCGGCTGACGCGTGGCTGCGGATTTCTTGGCAAAATTAACAGCGTT 360
Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 361 AAGCGTGTGGTGAGCAACCCCAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 420

QY 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAAsnHisThrLeuAspPheAsp 160
 Db 421 GGTCTGGAAGTTCTGTAATACGCTTATTATGATGCGAAATACACACTCTTGACTTCGAT 480
 QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 481 GCACCTGATTAACAGCGCTGAATGAAGCTCAGGCTGCGACGATGATGCTGTTCCATGGCTGC 540
 QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 Db 541 TCCCATTAACCCAAACCGGTATGACCTTACGCTGCGAACAATGCAACACTGGCACACTC 600
 QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlnPheAlaArgGly 220
 Db 601 TCCGTTTGAGAAGGCTGTTTACCGCTGTTTGAATTCGCTTACACGGGTTTGGCCGCTGGT 660
 QY 221 LeuGluGluAspAlaGlyLeuArgAlaPheAlaAlaMetHisGlyGluLeuIleVal 240
 Db 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCCGCGGCTATGCATAAAGAGCGTGTATT 720
 QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 Db 721 GCCAGTCTCTCTCTAAACAACTTGGCCCTGTACAACGAGCGGTGGCGCTTGTACTCTG 780
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 781 GTTGTGCGGACAGTGAACCGTTGATCGCGATTCAGCCAAATGAAGCGGCGATTGCG 840
 QY 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 Db 841 GCTAACTACTCTAAACCCACGACACACGCGCTTCTGTTGTTGCCACCATCTCGAGCAAC 900
 QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 901 GATGCGTTTACGCGGATTTGGGAACAAGAGCTGATGATGCGCGACGCTATTGAGCGT 960
 QY 321 MetArgGlnLeuPheValAsnThrLeuGlnLysGlyAlaAsnArgAspPheSerPhe 340
 Db 961 ATGCGTCACTGTTGTTGTCATACGCTGCAGGAAAGGCGCAACCGCGACTTCAGCTTT 1020
 QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1021 ATCATCAACAGAACCGCATGTTCTCTTCACTGCGCTGACAAAAGAACAAAGTGTGCGT 1080
 QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1081 CTGCGCGAAGAGTTTGGCGTATATGCGTTGCTTCTGCTGCGGTAAATGTGGCGCGGATG 1140
 QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 1141 ACACCAAGATAACATGGCTCCGCTGCGAAGCGATTGTGGCAGTCTG 1188

RESULT 5

US-09-801-563-87/c
 ; Sequence 87, Application US/09801563
 ; GENERAL INFORMATION:
 ; APPLICANT: Levy, Stuart, et. al.
 ; TITLE OF INVENTION: NMR COMPOSITIONS AND THEIR METHODS OF USE
 ; FILE REFERENCE: PKZ-043
 ; CURRENT APPLICATION NUMBER: US/09/801,563
 ; PRIORITY FILING DATE: 2001-03-08
 ; PRIOR APPLICATION NUMBER: 60/188,362
 ; PRIOR FILING DATE: 2000-03-10
 ; NUMBER OF SEQ ID NOS: 98
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 87
 ; LENGTH: 14759
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 US-09-801-563-87

Alignment Scores: 2.04e-312 Length: 14759
 Pred. No.:

Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 Gaps: 0
 DB: 33
 US-10-673-786A-2 (1-396) x US-09-801-563-87 (1-14759)
 QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 2755 ATGTTTGGAACATTAACCGCGCTCTCGCGACCGGATCTGGGCGTGGCGGATCTGTT 2696
 QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 Db 2695 CGTCCGATGAACGTCGCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 2636
 QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
 Db 2635 GGCAGAAACCCCGGTACTGACCGCGTGAAGAGCTGAACAGTATCTCTCGAAATGAA 2576
 QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyValGlyCysThrGlnGlu 80
 Db 2575 ACCACCAAAATTAACCTCGGCAATTCGCGCATCCCTGAATTTGGTCGCTGCACCTCAGNA 2516
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 2515 CTGCTGTTTGTAAAGGTAGCGCTGATCAATGACAAACGTCGCGACGACAGACT 2456
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 2455 CCGGGGGGCACTGGCGCACTACGCGTGGCTCCGATTTCTGGCAAAAAATACCAACGCTT 2396
 QY 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 Db 2395 AAGCGTGTGGGTGAGCAACCCAGCTGGCGCAACCAATAGAGAGGCTCTTTAACTCTGCA 2336
 QY 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAAsnHisThrLeuAspPheAsp 160
 Db 2335 GGTCTGGAAGTTCTGTAATACGCTTATATGATGCGGAATATCACACTCTTGACTTCGAT 2276
 QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 2275 GCACCTGATTAACAGCTGAATGAAGCTCAGGCTGCGACGATGCTGCTTCCATGGCTGC 2216
 QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 Db 2215 TGCCATAACCAACCGGTATCGACCTACGCTGGGAACAATGGCAACACTGGCAACATC 2156
 QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 2155 TCCGTTGGAAGAGCTGGTTACCGCTGTTGACCTTCGCTTACCAGGTTTGGCCGCTGGT 2096
 QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 Db 2095 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCCGCGCTATGCATAAAGAGCGTATTGTT 2036
 QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 Db 2035 GCCAGTTCTCTACTCTAAAAACCTTTGGCGCTGTACACGAGCGGTGGCGCTTGTACTCTG 1976
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 1975 GTTGTGCGGACAGTGAACCGTTGATCGCGATTCAGCCAAATGAAGCGGCGATTGCG 1916
 QY 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 Db 1915 GCTAACTACTCTAACCCACGACGACGCGCTTCTGTTGTTGGCCACCATCTCTGAGCAAC 1856
 QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 1855 GATGCGTTTACGCGATTTGGGAACAAGAGCTGATGATGCGCGACGCTATTTCAGCGT 1796
 QY 321 MetArgGlnLeuPheValAsnThrLeuGlnLysGlyAlaAsnArgAspPheSerPhe 340

```
|||||
Db 1795 ATGGCTCAGTTGTTGCTCAATACGCTGCAGGAAAGCGCAAAACCGCGACTTCAGCTTT 1736
Qy 341 lletleLysGlnAenGlyMetPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1735 ATCATCAACAGACGCGGATGTTCTCCTTCAGTGGCCTGACAAAAGAACAAGTCTGCGT 1676
Qy 361 LeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAenValAlaGlyMet 380
Db 1675 CTGGCGAAGAGTTGGCGTATATGCGGTTGCTTCTGCTCGCTAAATGTGGCGGGATG 1616
Qy 381 ThrProAenMetAlaProLeuGlyCysGluAlaValAlaValLeu 396
Db 1615 ACACAGATAAATGCGCTCCGCTGTGCGAAGCGATTGTGGCAGTGCTG 1568

RESULT 6
US-10-893-671-87/c
; Sequence 87, Application US/10893671
; GENERAL INFORMATION:
; APPLICANT: Levy, Stuart, et. al.
; TITLE OF INVENTION: NMR COMPOSITIONS AND THEIR METHODS OF USE
; FILE REFERENCE: PK2-043
; CURRENT APPLICATION NUMBER: US/10/893,671
; CURRENT FILING DATE: 2004-07-15
; PRIOR APPLICATION NUMBER: US/09/801,563
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/188,362
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 87
; TYPE: DNA
; LENGTH: 14759
; ORGANISM: Escherichia coli
US-10-893-671-87

Alignment Scores:
Pred. No.: 2,046-312 Length: 14759
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0
DB: 62

US-10-673-786A-2 (1-396) x US-10-893-671-87 (1-14759)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 2755 ATGTTTGAGAACATTACCGCGCTCCTGCCACCGCATTCGGGCTGGCGATCTGTTT 2696
Qy 21 ArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValTyLysAspGluThr 40
Db 2695 CGTGCGGATGAACGTCCTCCGCAAAATTAACCTCGGGATTGGTGCTATAAAGATGAGACG 2636
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyLeuLeuGluAenGlu 60
Db 2635 GCGAAACCCCGGTACTACCGAGCTGCCGCAAAATTAACCTCGGGATTGGTGCTATAAAGATGAGACG 2576
Qy 61 ThrThrLysAenTyLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 2575 ACCACCAAAATTAACCTCGGATTAACCGGCTCCCTGAAATTTGGTCGCTGACATCAGGAA 2516
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
Db 2515 CTGCTGTTTGAAGGTAGCGCTTATGATGACAAACGCTGCTCGCACGGCAGACT 2456
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 2455 CCGGGGGCGACTGGGGCCTACGCGTGGCTGCCGATTTCTGGCAAAAAATACACGCGTT 2396
Qy 121 LysArgValTropValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 2395 AAGCGTGTGGGTGAGCAACCCAGCTGGCGGCAACCAAGAGCGCTTTTAACCTCTGCA 2336
```

```
Qy 141 GlyLeuGluValArgGluTyAlaTyTyAspAlaGluAenHisThrLeuAspPheAsp 160
Db 2335 GGTCTGGAAGTTTCGTGAATACGCTTATATGATGCGGAAATACACCTCTTGACTTCGAT 2276
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 2275 GCACTGATTAACAGCTGAATGAAGCTCAGGCTGGCGACGTAGTCTGTTCCATGGCTGC 2216
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 2215 TGCCATAACCCACCGGTATCGACCTACGCTGGAACAATGGCAACAACCTGGCAACAATC 2156
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyGlnGlyPheAlaArgGly 220
Db 2155 TCCGTTGAGAAAGCTGTTTACCGCTGTTTACCTTCGCTTACCCAGGGTTTGGCCGCTGT 2096
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 2095 CTGGAAGAAGATGCTGAAGGACTGGCGGCTTTCGGGCTATGCAATAAAGAGCTGATGTT 2036
Qy 241 AlaSerSerTySerLysAenPheGlyLeuTyAenGluArgValGlyAlaCysThrLeu 260
Db 2035 GCCAGTTCCTACTCTAAAAACTTTGGCTCTGTACAAACGAGCGTGTGGCGCTTGACTCTG 1976
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 1975 GTTGCTGCCGACAGTGAACCGTTGATCGGCGCATTCAGCCAAATGAAAGCGCGCATTCGC 1916
Qy 281 AlaAenTySerAenProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
Db 1915 GCTAACTACTCTAACCCACCGACACGCGCTTCTGTTGTGCACCATCTCTGAGCAAC 1856
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 1855 GATGCGTTACGTGCGATTTCGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCGT 1796
Qy 321 MetArgGlnLeuPheValAenThrLeuGlnGlyLysGlyAlaAenArgAspPheSerPhe 340
Db 1795 ATGCGTCAGTTGTTGCTCAATACGCTCAGGAAAGGCGCAAAACCGCGACTTCAGCTTT 1736
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1735 ATCATCAACAGACGCGCATGTTCTCCTTCAGTGGCTTGACAAAGAACAGTCTGCGT 1676
Qy 361 LeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAenValAlaGlyMet 380
Db 1675 CTGGCGCAAGAGTTTGGCTATATGCGGTTGCTTCTGCTCGCTAAATGTGGCGGGATG 1616
Qy 381 ThrProAenMetAlaProLeuGlyCysGluAlaIleValAlaValLeu 396
Db 1615 ACACAGATAAATGCGCTCGCTGTGCGAAGCGATTGTGGCAGTGCTG 1568

RESULT 7
US-60-709-960-1/c
; Sequence 1, Application US/60709960
; GENERAL INFORMATION:
; APPLICANT: Frederick, Blattner
; TITLE OF INVENTION: REDUCED GENOME E. COLI
; FILE REFERENCE: 02730,0017,PZUS00
; CURRENT APPLICATION NUMBER: US/60/709,960
; CURRENT FILING DATE: 2005-08-20
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 4639675
; TYPE: DNA
; ORGANISM: Escherichia coli
US-60-709-960-1

Alignment Scores:
Pred. No.: 1,9e-307 Length: 4639675
Score: 2045.00 Matches: 396
```

Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 83 Gaps: 0

US-10-673-786A-2 (1-396) x US-60-709-960-1 (1-4639675)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 984932 ATGTTTGAACAATTAACCGCGCTCTCGCGACCGGATCTGGCGCTGGCGATCTGTT 984973

Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrIleAspGluThr 40
 Db 984872 CGTCCGATGAACGTCGCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 984813

Qy 41 GlyLysThrProValLeuThrSerValLysIleAlaGluGlnTyrLeuLeuGluAsnGlu 60
 Db 984812 GGCNAACCCCGGTACTACCGAGCGTGAAAGGCTGAACAGTATCTGCTCGAAATGAA 984753

Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 984752 ACCACAAAATTAACCTCGCATTTGACGCGATCCCTGATTTGGTGGCTGCACTCAGAA 984693

Qy 81 LeuLeuPheGlyLysGlySerAlaLeuLeuAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 984692 CTGCTGTTTGGTAAAGTAGCCCTGATCAATGACAAACGCTGCTCGCACGACAGACT 984633

Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 984632 CGGGGGGACCTGGCGCATACGCTGGCTGCGGATTTCTGGCAAAAATACACGCGTT 984573

Qy 121 LysArgValTropValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 Db 984572 AAGCGTGTGGGTGAGCAACCAAGCTGCGCGCAACCAATAGAGCGTCTTTAACTCTGCA 984513

Qy 141 GlyLeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 Db 984512 GGTCTGGAAGTTCGTGAATACGCTATTATGATGCGGAAATACACACTCTTGCATTCGAT 984453

Qy 161 AlaLeuLysAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 984452 GCACGTGATTAACAGCTGAATGAAGCTCAGGCTGCGAGCTAGTGTGTTTCCATGGCTGC 984393

Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
 Db 984392 TGCCATAACCCACCGGTATGACCTACGCTGGCAACATGCAACACTGGCACAACCTC 984333

Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 984332 TCCGTTGAGAAAGCTGTGTTACGCTGTTTGTACTTACCGGGTTTTCGCCGTTGGT 984273

Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMethHisLysGluLeuVal 240
 Db 984272 CTGGAAGAAGATGCTGAAGACTCGCGCTTTTCGGCGCTATGCAATAAGAGCTGATGTT 984213

Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 Db 984212 GCCAGTTCCTACTCTAAAAAATTTGGCCCTGTACAACGAGCGTGTGGCGCTTGTACTCTG 984153

Qy 261 ValAlaAlaAspSerGluThrValAlaAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 984152 GTTGCTGCCGACGAGTGAACCGTGTATCGCGCATTTAGCCCAATGAACGCGCGATTCOC 984093

Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 Db 984092 GCTAACTACTCTAACCCACGACACACGCGCTTCTGTTGTTGCCACCATCTCGAGCAAC 984033

Qy 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 984032 GATCGGTACGTCGATTTGGGAACAAGAGCTGACTGATATGCGCAGCGTATTACGCGT 983973

Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnLysGlyValAsnArgAspPheSerPhe 340

Db 983972 ATGCGTCAGTTGTTGCTCAATACGCTGAGGAAAGGCGCAACCGCAGCTTACAGCTTT 983913

Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 983912 ATCATCAACAGAACGCGCATGTTCTCTTCACTGGCCTGACAAAAGAACAGTGTGCGT 983853

Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 983852 CTGCGCAAGAGTTTGGCGTATATCGGTTCTTCTGTCGCGTAAATGTGCGCGGATG 983793

Qy 381 ThrProAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 983792 ACACAGATAACATGGCTCCGCTGCGAAGCATTTGGCAGTCTG 983745

RESULT 8
 US-60-643-717-112
 ; Sequence 112, Application US/60643717
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark S.
 ; TITLE OF INVENTION: Genes and Uses for Plant Improvement
 ; FILE REFERENCE: 38-21(53629)A
 ; CURRENT APPLICATION NUMBER: US/60/643,717
 ; CURRENT FILING DATE: 2005-01-12
 ; NUMBER OF SEQ ID NOS: 19247
 ; SEQ ID NO 112
 ; LENGTH: 1576
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli K12
 US-60-643-717-112

Alignment Scores:
 Pred. No.: 2,186-313 Length: 1576
 Score: 2039.00 Matches: 395
 Percent Similarity: 99.7% Conservatives: 0
 Best Local Similarity: 99.7% Mismatches: 1
 Query Match: 99.7% Indels: 0
 DB: 82 Gaps: 0

US-10-673-786A-2 (1-396) x US-60-643-717-112 (1-1576)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 298 ATGATTGAGAACATTTACCGCGCTCTCGCGACCGGATTTCTGGCGCTGGCGATCTGTT 357

Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrIleAspGluThr 40
 Db 358 CGTCCGATGAACGTCGCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 417

Qy 41 GlyLysThrProValLeuThrSerValLysIleAlaGluGlnTyrLeuLeuGluAsnGlu 60
 Db 418 GGCNAACCCCGGTACTGACCGCGTGAAAGGCTGAACAGTATCTGCTCGAAATGAA 477

Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 478 ACCACAAAATTAACCTCGCATTTGACGCGATCCCTGAAATTTGGTGGCTGCACTCAGAA 537

Qy 81 LeuLeuPheGlyLysGlySerAlaLeuLysAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 538 CTGCTGTTTGGTAAAGTAGGAGCGCTGATCAATGACAAACGCTGCTCGCACGCGACACT 597

Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 598 CGGGGGGACCTGGCGCACTACGCGTGGCTGCGGATTTCTGGCAAAAATACACGCGTT 657

Qy 121 LysArgValTropValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 Db 658 AAGCGTGTGGGTGAGCAACCCAGCTGCGCAACCAATAGAGCGTCTTTAACTCTGCA 717

Qy 141 GlyLeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 Db 718 GGTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAATACACACTCTTGGACTTCGAT 777

Qy 161 AlaLeuLysAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180


```

Db 778 GCACGTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGAGTAGTGCTGTTCCATGGCTGC 837
Qy 181 CysHisAsnProThrGlyLeuAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 838 TGCCATAACCCCAACCGGTATCGACCTACGCTGGACCAATGGCAACATGGCAACACTGC 897
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTrpGlnGlyPheAlaArgGly 220
Db 898 TCCGTTGAGAAAGCGTGTATCCGCTGTTGACCTTCGCTTACCAAGGGTTTGGCCCGTGT 957
Qy 221 LeuGluGlnAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuVal 240
Db 958 CTGGAAGAAGATGCTGAAGAGACTGGCGCTTTGCGCGCTATGCAATAAAGAGACTGATT 1017
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 1018 GCCAGTTCCTACTCTAANAACCTTGGCCCTGTACACAGAGCGTGTGGCGCTTGACTCTG 1077
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 1078 GTTGCTGCCGACAGTGAACCGCTTGATCGCGCAATTCAGCCAAATGAAGCGCGATTGCG 1137
Qy 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 1138 GCTAACTACTCTAAACCCACACAGCACAGCGCGCTTCTGTGTGGCACCATCTCAGCAAC 1197
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 1198 GATGCGTTACGTGCGATTTGGGACACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 1257
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 1258 ATGCGTCAGTTGTTGCTCAATACGCTGCAGGAAAGGCGCAACCCGCGACTTCAGCTTT 1317
Qy 341 IleIleLysGlnAsnGlyWetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1318 ATCATCAACAGAACGCGCATGTTCTCTTCAGTGGCCCTGCACAAAGAACAAAGTCTCGGT 1377
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1378 CTGCGCGAAGAGTTGGCGTATATGCGGTTGCTTCTGCTCGGTAAATGTGGCGCGGATG 1437
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1438 ACACCAAGATAACATGGCTCCGCTGTGCGAAGCGATTGTGGCAGTGCTG 1485

```

RESULT 9

```

US-11-114-922-75
; Sequence 75, Application US/11114922
; GENERAL INFORMATION:
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; CURRENT FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 75
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
US-11-114-922-75

```

Alignment Scores: 3.87e-309 Length: 1191
 Pred. No.: 2011.00 Matches: 390
 Score:

```

Percent Similarity: 99.2% Conservative: 3
Best Local Similarity: 98.5% Mismatches: 3
Query Match: 98.3% Indels: 0
DB: 70 Gaps: 0
US-10-673-786A-2 (1-396) x US-11-114-922-75 (1-1191)
Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTTTGAGAACATTAACCGCGCTCTCTGCGACCCGATTCCTGGCCCTGGCCGATCTGTTT 60
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyLysGlyValTyrLysAspGluThr 40
Db 61 CGTCCCGATGAACGCTCCCGCAAAATTAACCTCGGATTTGGTCTCTATTACGATGAGACG 120
Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 121 GGCAAAATCCCGGTACTGACCGCGTGAAGAGGCTGAACAGTATCTGCTCGAAAAATGAA 180
Qy 61 ThrThrLysAsnTyrLeuGlyLysIleAspGlyLysProGluPheGlyArgCysThrGlnGlu 80
Db 181 ACCACCAAACTTTACCTCGGCATTTGACGGCATCTCCCTGAATTTGGTCTGCTGACCTCAG 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCTCGCACGGCACAGACT 300
Qy 101 ProGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 301 CCGGGGGGCTCTGGCGCACTACGCGTGGCTGCCGATTTCTTGCGCAAAAAATACCGCGTT 360
Qy 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 361 AAGCGTGTGGGTGAGCAACCCCAAGCTGCGCGCAACCATTAAGAGCGCTTTTAACCTCTGCA 420
Qy 141 GlyLeuGlnValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 421 GGCTGGAAGTTGCGTAATACGCTTATATGATGCGGAAATACACTCTTTCATCTTGAT 480
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 481 GCACGTATTAACAGCTGATGAAGCTCAGCGTGGCGACGTAGTGTCTCCATGGCTGC 540
Qy 181 CysHisAsnProThrGlyLysIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 541 TGCCATAACCCCAACCGGTATCGACCTTACGCCCTACGCTGGAACAATGGCAACACTTGC 600
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 601 TCCGTTGAGAAAGCGTGGTTACCGCTGTTTGACTTTCGCTTACCGGGTTTGGCCCGTGT 660
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTGCGCGCTATGCATAAAGAGCTGATTGT 720
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 721 GCCAGTTCTTACTCTAAAAAATTTGGCGCTGTACAAACGAGCGTGTGGCGCTTGACTCTG 780
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 781 GTTGCTGCCGACAGTGAACCGTGTGATCGCGCATTCAGCCAAATGAAGCGCGGATTCGC 840
Qy 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 841 GCTAACTACTCTTCCCGCACGACGCGCTTCTGTTGTTGCGCACCATCTCTGAGCAAC 900
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 901 GATGCGTTACGTGCGATTTGGGAACAAGAGCTGATGATGCGCCGACGCTATTTCAGCGT 960
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340

```


Db 961 ATGGCTGAGTTTCTGCTCAATACGCTGACGAGAAAGGCGCAACCGCGACTTTCAGCTTT 1020
 Qy 341 llelleLysGluAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1021 ATCATCAACAGACAGCGCATGTTCTCTTCACTGCGCTGACAAAGACAACTGCTGGT 1080
 Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1081 CTGCGCAAGAGTTTGGCGTATATGCGTTGCTTCTGCTGCGTAAATGTCGCGGATG 1140
 Qy 381 ThrProAspAsnMetAlaProLeuGluCysGluAlaIleValAlaValLeu 396
 Db 1141 ACACAGATAACATGCTGCTGCGGAGCGATTTGTGCGAGTGTGCTG 1188

RESULT 10

US-10-446-203-4902
 ; Sequence 4902, Application US/10446203
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/10/446,203
 ; CURRENT FILING DATE: 2003-05-27
 ; PRIOR APPLICATION NUMBER: US/09/489,039
 ; PRIOR FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 4902
 ; LENGTH: 1209
 ; TYPE: DNA
 ; ORGANISM: Klebsiella pneumoniae
 US-10-446-203-4902

Alignment Scores:

Pred. No.: 2,7e-296 Length: 1209
 Score: 1931.00 Matches: 370
 Percent Similarity: 96.5% Conservativity: 12
 Best Local Similarity: 93.4% Mismatches: 14
 Query Match: 94.4% Indels: 0
 DB: 51 Gaps: 0

US-10-673-786a-2 (1-396) x US-10-446-203-4902 (1-1209)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 19 ATGTTTGAGAACATTACCGCCGCCCGCCAGCCGACGATTTTAGGTCTGCGCGATCTGTT 78
 Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 Db 79 CTGCGGATGACCGCCCTGAAATAATTACCTCGGAATGGGTGTTTACAGATGAACACC 138
 Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuGluAsnGlu 60
 Db 139 GGTAAACCGCGTTCTGACACGAGCTCAAAAAGCAGACGACTATCTGCTGGAATGAA 198
 Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 199 ACGACTAAACATATCTGGGCATCATGATGTTTCTGAAATTTGGTCTGCTGACCCAGGAG 258
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 259 CTGCTGTTTCGGTAAGGCAATCGGATTCGCTGATTAACGCGCCGACCGCGAGACA 318
 Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 319 CCCGCGGTACCGGTGCGTGTGCTGCGCGAGACTTCTCTGCCCAAAACACACCGACGTG 378
 Qy 121 LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 Db 379 AAACGTGTGGGTAAAGTAATCCGAGTGTCCGCAACCATAAAGGCTATTACCTCTGCC 438

Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 Db 439 GGGCTGGAAGTCGCGCAATACGCATACGACGCGGCTAACACCGCTGCGACTTTGAT 498
 Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValLeuPheHisGlyCys 180
 Db 499 GGTCTGCTGGCGAGCTGAACGAAGCCAGCGCGCGCGTGTACTGTTCACCGCTGC 558
 Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
 Db 559 TGCCACACCGCAGCGGTATCGATCCGACGCTCGATCAGTGGCAGCAGCTGCGCAGCTG 618
 Qy 201 SerValGluLysGlyTyrProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 619 TCGGTGGAAGAGCTGCTACCGCTGTTTCGATTTTCGCTACACGAGGCTTCGCGCGGT 678
 Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 Db 679 CTGGAAGAAGATGCTGAAGGGCTACCGCTTTTGTGCTCTACATAAAGAGCTGCTGGTC 738
 Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 Db 739 GCCAGCTCTACTCGAAAAACTTTGGCCTGTACAACGAGCGGCTCGCGCGCTGCACCTG 798
 Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 799 GTCGCGCGGATCAGGAGACTGTAGACCGCGCTTCAGTCAGATGAAGTCGGTGTATCCGC 858
 Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 Db 859 CCAACTACTCGAACCCCGCTCGCATGCGCCTCCGTGGTGGCGCCACCATTCCTGAGCAAC 918
 Qy 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 919 GATCGCTACCGGCATCTGGAGCAGGAATGACCGATATGCGCAGCGCATCCAGCGT 978
 Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
 Db 979 ATGGCTGCTGTTTCGTCAATACCTGCGAGAGAAAGCGCGAGCGGACTTCAGCTTT 1038
 Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1039 ATCAGCCAGCAGAACCGCATGTTCTCATTCAGCGCGCTGACTAAAGACAGCGTGTGCGC 1098
 Qy 361 LeuArgGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1099 CTGGTGAAGAGTTGCGCATCTATCGGTAGTCTCCGAGCGTATCAAGTGGCGGATG 1158
 Qy 381 ThrProAspAsnMetAlaProLeuGluCysGluAlaIleValAlaValLeu 396
 Db 1159 ACGCTGACATATGGCGCGCTGTGCGAAGCCATCGTCCGCTACTG 1206

RESULT 11

US-09-252-691-1606
 ; Sequence 1606, Application US/09252691B
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith G. Weinstein et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
 ; FILE REFERENCE: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.135
 ; CURRENT APPLICATION NUMBER: US/09/252,691B
 ; CURRENT FILING DATE: 1999-02-18
 ; NUMBER OF SEQ ID NOS: 11324
 ; SEQ ID NO 1606
 ; LENGTH: 1209
 ; TYPE: DNA
 ; ORGANISM: Enterobacter cloacae
 US-09-252-691-1606

Alignment Scores:
 Pred. No.: 3.91e-296 Length: 1209
 Score: 1930.00 Matches: 370
 Percent Similarity: 97.0% Conservativity: 14

Best Local Similarity: 93.4% Mismatches: 12
 Query Match: 94.4% Indels: 0
 DB: 22 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-252-691C-1606 (1-1209)

Qy 1 MetPheGluAenIleThrAlaAProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
 Db 19 ATGTTTGAGAACATTAATCTGCGCGCTCTGCGGCAATTTCTGGGCGCTGGCGATCTGTTT 78
 Qy 21 ArgAlaAaspGluAArgProGlyLysIleAenLeuGlyLysGlyValTyrLysAaspGluThr 40
 Db 79 CGCGCCGACGACCCCTCGGAAATCAACTGGGTATTTGGTGTATATAAGATGAAC 138
 Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
 Db 139 GGCAAACTCGGTACTGACCGCGTCAAAAAGCTGAGCAGTATCTGCTGGAACGAA 198
 Qy 61 ThrThrLysAenTyrLeuGlyLysIleAaspGlyLysProGluPheGlyArgCysThrGlnGlu 80
 Db 199 ACCACCAAACTACCTCGGTATTTGATGATCTGCGGCAATTTCTGGGCGCTGGCGAG 258
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuLysAenLysAaspLysArgAlaAargThrAlaGlnThr 100
 Db 259 CTGCTGTTTCGCAAGGACGACCATTTGTGAGCGAAACCGCGCGACGCGCGACGACC 318
 Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAaspPheLeuAlaLysAenThrSerVal 120
 Db 319 CCAGCGGTACCGGCACTCGGCGTGGCGACATTTCTGGCGAAACACATCTCTGTG 378
 Qy 121 LysArgValTyrValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
 Db 379 AAGCGTGTGGTAAAGCAATCAAGCTGGCGCAACCATTAAGAGCGTGTAAATCTCGC 438
 Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAaspPheAasp 160
 Db 439 GGTCTGGAAGTGCCTGATACCGCTACTACGACGACGCGCGCGCTGGATTTTCGAC 498
 Qy 161 AlaLeuLysAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCys 180
 Db 499 GGCTGTGCGCAGCTGCGGAGGCGCGCGCGGCGATGTTGGTGTTCACGCGTGC 558
 Qy 181 CysHisAenProThrGlyLysAaspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
 Db 559 TGCCATAACCAACCGGTATCGATCTACGCTTACGAGTGGGAACAGCTGGCTAAGCTG 618
 Qy 201 SerValGluLysGlyTyrLeuProLeuPheAaspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 619 TCCGTTGAAAAGGCTGCGCTCGCTGTTGACTTTCGCTTACCGGCGCTTTGCGCGTGT 678
 Qy 221 LeuGluGluAaspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
 Db 679 CTGGAAGAAGATGAGAAAGTCTGCGCGCATTCGCGCGCTGATCAAGATTTGATGTC 738
 Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyAlaCysThrLeu 260
 Db 739 GCAAGCTCTATTCCAAGAACTTTGGTCTGTATTAACGAGCGTGTGCGCGCTGTACGCTG 798
 Qy 261 ValAlaAaspSerGluThrValAaspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 799 GTGGTGTGTCGAGCAACCGTGCATTCGCGCATTTACGCCAGATGAAGTTCGCTGATCCGC 858
 Qy 281 AlaAenTyrSerAenProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
 Db 859 GCTAACTACTCCAAACCCGCGCACCGTGGTGTGCGGCGCATCTGAGCAAC 918
 Qy 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAaspMetArgGlnArgIleGlnArg 320
 Db 919 GATGCGTGTGCGGCACTCTGGGAGCAAGAGCTGAACGATATGCGCTCAGCGCATTTACGCGC 978
 Qy 321 MetArgGlnLeuPheValAenThrLeuGlnGlyLysGlyAlaAenArgAaspPheSerPhe 340
 Db 979 ATGCGTCTGCTGTTTGTGAACACGCTGCGTGAAGAGCGCTGACCGGTGACTTCAGCTTC 1038

Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1039 ATCATCAAGCAGAACGCGCATTTCTCTTCAGCGCGCTGACCAAGAGCAGGCTCTCGC 1098
 Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenValAlaGlyMet 380
 Db 1099 CTGCGTGAAGAGTTTGGCGGTATGCGGTGCGATCTGCGCGCGTGAACGTTGCAAGGTATG 1158
 Qy 381 ThrProAaspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 1159 ACAGCTGACAACTGCGCGCTGTGCGAAGCCATTGTCGCGGTGCTG 1206

RESULT 12
 US-09-252-691C-1606
 ; Sequence 1606, Application US/09252691C
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith G. Weinstein et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
 ; TITLE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.135
 ; CURRENT APPLICATION NUMBER: US/09/252.691C
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,145
 ; PRIOR FILING DATE: 1998-07-24
 ; PRIOR APPLICATION NUMBER: US 60/074,787
 ; PRIOR FILING DATE: 1998-02-18
 ; NUMBER OF SEQ ID NOS: 11326
 ; SEQ ID NO 1606
 ; LENGTH: 1209
 ; TYPE: DNA
 ; ORGANISM: Enterobacter cloacae
 US-09-252-691C-1606

Alignment Scores:
 Pred. No.: 3,91e-296 Length: 1209
 Score: 1930.00 Matches: 370
 Percent Similarity: 97.0% Conservative: 14
 Best Local Similarity: 93.4% Mismatches: 12
 Query Match: 94.4% Indels: 0
 DB: 22 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-252-691C-1606 (1-1209)

Qy 1 MetPheGluAenIleThrAlaAProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
 Db 19 ATGTTTGAGAACATTAATCTGCGCGCTCTGCGGCAATTTCTGGGCGCTGGCGATCTGTTT 78
 Qy 21 ArgAlaAaspGluAArgProGlyLysIleAenLeuGlyLysGlyValTyrLysAaspGluThr 40
 Db 79 CGCGCCGACGACCCCTCGGAAATCAACTGGGTATTTGGTGTATATAAGATGAAC 138
 Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
 Db 139 GGCAAACTCGGTACTGACCGCGTCAAAAAGCTGAGCAGTATCTGCTGGAACGAA 198
 Qy 61 ThrThrLysAenTyrLeuGlyLysIleAaspGlyLysProGluPheGlyArgCysThrGlnGlu 80
 Db 199 ACCACCAAACTACCTCGGTATTTGATGATCTGCGGCAATTTCTGGGCGCTGGCGAG 258
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuLysAenLysAaspLysArgAlaAargThrAlaGlnThr 100
 Db 259 CTGCTGTTTCGCAAGGACGACCATTTGTGAGCGAAACCGCGCGACGCGCGACGACC 318
 Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAaspPheLeuAlaLysAenThrSerVal 120
 Db 319 CCAGCGGTACCGGCACTCGGCGTGGCGACATTTCTGGCGAAACACATCTCTGTG 378
 Qy 121 LysArgValTyrValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
 Db 379 AAGCGTGTGGTAAAGCAATCAAGCTGGCGCAACCATTAAGAGCGTGTAAATCTCGC 438
 Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAaspPheAasp 160


```
Db 919 GATCGCTGCGCGCATCTGGGAGCAAGAGCTGAACGATATGCTGCTCAGCGCATTCAGCGC 978
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 979 ATGCGCTGCTGTTTGTGAACACGCTGGCTGAGAAAGCGCTGACCGTGACTTCAGCTTC 1038
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1039 ATCATCAAGCAGAACGGCATGTTCTCTTCAGCGCCTGACCAAGAGCAGGTGCTGCGC 1098
Qy 361 LeuArgGluGluPheGlyValThrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1099 CTGCGTGAAGAGTTGCGGTGATGCGCGTTCATCTGCGCGCGTGAACGTTGCAAGTATG 1158
Qy 381 ThrProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1159 ACGCCTGACAACATGGCGCGCTGTGCGAAGCCATTGTCGGCGGTGCTG 1206

RESULT 14
US-09-252-691C-1616/c
; Sequence 1616, Application US/09252691B
; GENERAL INFORMATION:
; APPLICANT: Keith G. Weinstock et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
; FILE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.135
; CURRENT APPLICATION NUMBER: US/09/252,691B
; NUMBER OF SEQ ID NOS: 11324
; SEQ ID NO 1616
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Enterobacter cloacae
US-09-252-691-1616

Alignment Scores:
Pred. No.: 3,97e-296 Length: 1218
Score: 1930.00 Matches: 370
Percent Similarity: 97.0% Conservative: 14
Best Local Similarity: 93.4% Mismatches: 12
Query Match: 94.4% Indels: 0
DB: 22 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-252-691-1616 (1-1218)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1200 ATGTTTGAGAACATTAATCTGCGCGCTCTGCGGACCAATTTCTGGCGCTGGCGATCTGTTT 1141
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValThrLysAspGluThr 40
Db 1140 CGCGCCGACGACCGCCCTGGAAAATCACTGCTGGGTATTTGGTGTATATAAGATGAAC 1081
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnThrLeuLeuGluAenGlu 60
Db 1080 GGCAAACTCCGGTACTGACCGAGCTCAAAAAGCTGAGCAGTATCTGCTGGAACGAA 1021
Qy 61 ThrThrLysAenThrLeuGlyLysAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 1020 ACCACCAAACTACCTCGGTATGATGGTATCTCTGAAATTTGTCGCTGCACCGAG 961
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 960 CTGCTGTTTCGCAAGGAGGAGCACCATTGTGAGCGAAAACCGCGCGGCGGCGGAGACC 901
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 900 CCAGCGGTACCGCGCATCTGCGGTGGCGCAGATTTCTTCTGGCGAAAACACATCTCTGTG 841
Qy 121 LysArgValThrValSerAsnProSerThrProAsnHisLysSerValPheAsnSerAla 140
Db 840 AAGCGTGTGTGGTGAAGCAATCCAAGCTGGCGCAACCAATAAGAGCGGTGTTTAATCTGCG 781
```

```
Qy 141 GlyLeuGluValArgGluThrAlaThrTyrAspAlaGluAenHisThrLeuAspPheAsp 160
Db 780 GGTCTGGAAGTGGGTGAATACGCTACTACGAGCGACGCCACCGCTGGATTCGAC 721
Qy 161 AlaLeuIleAsnSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 720 GGCCTGTGCGCCAGCTGAGCGAAGCCAGCGGCGATGTGTGCTGTTCACCGGTGTC 661
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnThrLeuAlaGlnLeu 200
Db 660 TGCCATAACCCACCGGTATCGATCTACGCTTGAGCAGTGGGAACAGCTGGCTAAGCTG 601
Qy 201 SerValGluLysGlyThrLeuProLeuPheAspPheAlaThrGlnGlyPheAlaArgGly 220
Db 600 TCGCTTGAATAAGCGCTGGCTGCGCTGTTTGTACTTCGCTTACCAGGCGTTTGGCCGTGGT 541
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMethHisLysGluLeuIleVal 240
Db 540 CTGGAAGAAGATGAGAGGTCTGCGCGCATTCGCCCGCTGCATCAAGAGTTGATAGTC 481
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 480 GCAAGCTCTCTATTCGAAGAACTTTGGTCTGTATAACGAGCGTGTGCGCGCTGTACGCTG 421
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaArg 280
Db 420 GTGCTGCTGACGAAGCAACCGCTCGATCGCGCATTCAGCCAGATGAAGTCCGTGATCCGC 361
Qy 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 360 GCTAACTACTCCAACCCACCGCACAGCGTGTGTCGCGCAGATCTCTGAGCAAC 301
Qy 301 AspAlaLeuArgAlaIleThrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 300 GATGCGCTGCGCGCATCTGGGAGCAAGAGCTGAACATATGCTGACGCGCATTCAGCGC 241
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 240 ATGCGTCTGCTGTTTGTGAACACGCTGGCTGCGTGAAGAGCGCTGACCGCTGACTTCAGCTTC 181
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 180 ATCATCAAGCAGAACCGCATGTTCTCTCTGCGCGCTGACCAAGAGCAGGTGCTGCGC 121
Qy 361 LeuArgGluGluPheGlyValThrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 120 CTGCGTGAAGAGTTTGGCGTGTATGCGGTGATCTGCGCGCGTGAACGTTGCAAGTATG 61
Qy 381 ThrProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 60 ACGCTGACAACATGGCGCGCTGTGGAAGCCATTGTCGCGGTGCTG 13
```

```
RESULT 15
US-09-252-691C-1616/c
; Sequence 1616, Application US/09252691B
; GENERAL INFORMATION:
; APPLICANT: Keith G. Weinstock et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
; FILE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.135
; CURRENT APPLICATION NUMBER: US/09/252,691C
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/094,145
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: US 60/074,787
; PRIOR FILING DATE: 1998-02-18
; NUMBER OF SEQ ID NOS: 11326
; SEQ ID NO 1616
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Enterobacter cloacae
US-09-252-691C-1616
```

Alignment Scores:

Pred. No.: 3,97e-296 Length: 1218
Score: 1930.00 Matches: 370
Percent Similarity: 97.0% Conservative: 14
Best Local Similarity: 93.4% Mismatches: 12
Query Match: 94.4% Indels: 0
DB: 22 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-252-691C-1616 (1-1218)

```
Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGluValLeuAlaAspLeuPhe 20
Db 1200 ATGTTTGAGAACATTACTCGCGCTCTCTGCCACCAATCTGGGCTCGCCGATCTGTTT 1141

Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGluGlyValIleGlyValIleAspGluThr 40
Db 1140 CGCGCCGACGCGCCCTGGAAAATCAACCTGGGTATTGGTGTATATAAGATGAAC 1081

Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 1080 GGCAAACTCCGGTACTGACGAGCTCAAAAAGCTGAGCAGTATCTCTGCAAAACGAA 1021

Qy 61 ThrThrLysAsnTyrLeuGlyLysIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 1020 ACCACCAAACTACTCGGTATTGATGGTATCCCTGAATTTGGTCTGACCCAGGAG 961

Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 960 CTGCTGTTCCGGAAGGAGGAGCACCATTGTGAGCGAAAACCGGCCCGCCGCGCAGACC 901

Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 900 CCAGCGGTACCGCGGCACCTGCGCTGCGGCAGATTTCTTGGCGAAAACACATCTCTGTG 841

Qy 121 LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 840 AAGCGTGTGGGTGAAGCAATCAACAGCTGGCGGACCAACCAAGAGCGTGTTTAATCTCGG 781

Qy 141 GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 780 GGTCTGGAAGTGGTGAATACGCTTACGAGCGGCGCAGCCGCGCTGGATTTTCGAC 721

Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 720 GGCTGTCTGGCAGGCTGAGCGAAGCCAGCGGCGGATGTGTCTGTCTCCAGCGCTGC 661

Qy 181 CysHisAsnProThrGlyLysAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 660 TGCCATAACCCAAACCGGTATCGATCTCTACGCTTGAGCAGTGGGAACAGCTGGCTAAGCTG 601

Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 600 TCCGTTGAAAAGGTGGCTGCGCTGTTGACTTTCGCTTACCGAGGCTTTGGCCCGTGT 541

Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 540 CTGGAGAGAGATGCAAGAGGCTGCGCGCATTCGCGCGGTGCATCAAGAGTGTATAGTC 481

Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 480 GCAAGCTCTATTCCCAAGAACTTTGGTCTGTATACGAGCGGTGTCGGCGCTGTACGCTG 421

Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 420 GTGGCTGTCTGCAAGAGCAACCGCTCGATTCGCGCATTCAGCCAGATGAAGTCCGTGATCCGC 361

Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 360 GCTAACTACTCCAAACCCCGGCAACCGGTGGCTCTGTGGTCCGACCATCTCTGAGCAAC 301

Qy 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 300 GATGCGCTGCGGCCATCTGGGAGCAAGAGCTGAACGATATGCGTCAAGCATTCAGCGC 241
```

Search completed: March 15, 2006, 17:48:21

Job time : 8035 secs

```
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 240 ATGCGTCTGCTGTTTGTGAACACACGCTGGCTGAGAAAGCGCTGACCGTGAAGTTCAGCTTC 181

Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 180 ATCATCAAGCAGAACCGCATGTTCTCTTCAGCGGCTGACCAAGAGCAGGTGCTGCGC 121

Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 120 CTGCGTGAAGAGTTTGGCGTGTATGCCGTTCATCTGSCCGGTGAACGTTGACAGTATG 61

Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 60 ACGCTGACAAACATGGCGCGCTGTCGAAGCCATTGTTCGCGGTGCTG 13
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 15:45:30 ; Search time 410 Seconds
(without alignments)
2682.737 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045
Sequence: 1 MFENITAPADPILGLADLF.....VAGMTPDNNAFLCAIVAVL 396

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 6590814 seqs, 1388790629 residues

Total number of hits satisfying chosen parameters: 13181628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/abse/ABSSWEB_spool/US10673786/runat_14032006_135618_22781/app_query.fasta_1
-DB=Pending_Patents_NA_New -QFMT=fastap -SUFFIX=rnpn -MINMATCH=0.1 -LOOPECL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abes03h -USER=US10673786 @CGN 1 1 70 @runat_14032006_135618_22781
-NCPU=6 -ICPU=3 -NO MAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents_NA_New.*

```
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US12_NEW_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US13_NEW_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US14_NEW_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US15_NEW_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US16_NEW_COMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2045	100.0	1191	10	US-11-302-262-19
2	2045	100.0	1241	1	PCT-US05-47111-4
3	2045	100.0	1241	12	US-60-752-355-4
4	2039	99.7	1576	1	PCT-US06-00964-112
5	2039	99.7	1576	8	US-11-330-403-112
6	1310	64.1	2242716	7	US-10-018-4708-1

C

7	913	44.6	1230	1	PCT-US05-47111-133	Sequence 133, App
8	913	44.6	1230	12	US-60-752-355-133	Sequence 133, App
9	909	44.4	1224	1	PCT-US05-47111-446	Sequence 446, App
10	909	44.4	1224	12	US-60-752-355-446	Sequence 446, App
11	885	43.3	1668	7	US-10-953-349-16412	Sequence 16412, A
12	865	42.3	1885	12	US-60-771-906-33	Sequence 33, Appl
13	814	39.8	1218	10	US-11-299-286-563	Sequence 563, App
14	806	39.4	1301	7	US-10-953-349-11025	Sequence 11025, A
15	794	38.8	2339	8	US-11-329-121-902	Sequence 902, App
16	794	38.8	2339	12	US-60-741-050-148	Sequence 148, App
17	794	38.8	2339	12	US-60-763-366-696	Sequence 696, App
18	794	38.8	2417	8	US-11-329-121-903	Sequence 903, App
19	794	38.8	2417	12	US-60-763-366-697	Sequence 697, App
20	794	38.8	2465	8	US-11-329-121-901	Sequence 901, App
21	794	38.8	2465	12	US-60-763-366-695	Sequence 695, App
22	792	38.7	2411	10	US-11-266-748A-25190	Sequence 25190, A
23	783	38.3	2005	10	US-11-266-748A-184254	Sequence 184254, A
24	783	38.3	2005	10	US-11-266-748A-191752	Sequence 191752, A
25	759	37.1	2365	10	US-11-266-748A-68736	Sequence 68736, A
26	759	37.1	2365	10	US-11-266-748A-121547	Sequence 121547, A
27	758	37.1	1293	10	US-11-299-286-769	Sequence 769, App
28	758	37.1	1612	7	US-10-953-349-3641	Sequence 3641, App
29	736	36.0	1937	10	US-11-224-525-639	Sequence 639, App
30	736	36.0	1937	10	US-11-224-525-639	Sequence 639, App
31	735.5	36.0	1299	1	PCT-US05-47111-2	Sequence 2, Appli
32	735.5	36.0	1299	12	US-60-752-355-2	Sequence 2, Appli
33	726.5	35.5	1230025	8	US-11-366-942-1	Sequence 1, Appli
34	652.5	31.9	1373	10	US-11-266-748A-68993	Sequence 68993, A
35	652.5	31.9	1373	10	US-11-266-748A-105184	Sequence 105184, A
36	652.5	31.9	1373	10	US-11-266-748A-121804	Sequence 121804, A
37	606	29.6	2777	10	US-11-266-748A-68992	Sequence 68992, A
38	606	29.6	2777	10	US-11-266-748A-105183	Sequence 105183, A
39	606	29.6	2777	10	US-11-266-748A-121803	Sequence 121803, A
40	598.5	29.3	1356	1	PCT-US05-47111-3	Sequence 3, Appli
41	598.5	29.3	1356	12	US-60-752-355-3	Sequence 3, Appli
42	594.5	29.1	1045	10	US-11-266-748A-9142	Sequence 9142, Ap
43	594.5	29.1	1045	10	US-11-266-748A-64005	Sequence 64005, A
44	594.5	29.1	1045	10	US-11-266-748A-66837	Sequence 66837, A
45	494	24.2	801	8	US-11-360-355-48767	Sequence 48767, A

ALIGNMENTS

RESULT 1

```
US-11-302-262-19
; Sequence 19, Application US/11302262
; GENERAL INFORMATION:
; APPLICANT: Ajinomoto Co., Inc.
; TITLE OF INVENTION: A method of improving growth and yield of plants under reduced
; FILE REFERENCE: OP05262
; CURRENT APPLICATION NUMBER: US/11/302,262
; CURRENT FILING DATE: 2005-12-14
; PRIOR APPLICATION NUMBER: JP 2003-198559
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1188)
US-11-302-262-19
```

Alignment Scores:
Pred. No.: 2e-237 Length: 1191
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 0

US-10-673-786A-2 (1-396) x US-11-302-262-19 (1-1191)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTTTGAGAACATTACCGCGCTCTGCGGACCGGATCTGGCGCTGGCGATCTGTTT 60
Qy 21 ArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValThrLysAspGluThr 40
Db 61 CGTCCGATGAACGTCCTGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 120
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 121 GGCACAAACCCGGTACTACCGCGCTGAAAGGCTGAACAGTATCTGCTCGAAATGAA 180
Qy 61 ThrThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 181 ACCACCAAAATTTACCTCGGATTTGACGGCATCTCTGAAATTTGGTGGCTGCATCAGGAA 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
Db 241 CTGCTGTTTGGTAAAGTAGCGCTGATCAATGACAAACGCTGTCGACGCGCACAGACT 300
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 301 CCGGGGGGCACCTGGCGCACTACGGCTGGCTGCCGATTTCTCGGCAAAATTAACAGCGTT 360
Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 361 AAGCGTGTGGTGGTAGGACCCAGCTGGCGGACCCATAGAGCGTCTTTAACTCTGCA 420
Qy 141 GlyLeuGluValArgGluTyrAlaTyrArgAlaGluAenHisThrLeuAspPheAsp 160
Db 421 GGTCTGGAAGTTCTGGTAATACGCTTATATGATCGGAAATCAACACTCTTTGACTTCGAT 480
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 481 GCACGTGATTAACAGCTGAAGCTCAGGCTGGCGAGTAGTGTCTGTTCCATGGCTGC 540
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 541 TGCCATACCCCAACCGGTATCGACCTACGCTGGACAAATGGCAAACTGGCAAACTC 600
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 601 TCCGTGTAGAAGGCTGTTACCGCTGTTCACCTTCGCTTACCGAGGTTTTCGCCGTGT 660
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCGCGGCTATGTCATAAAGAGCTGATTGTT 720
Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyValaCysThrLeu 260
Db 721 GCCAGTTCCTACTCTAAAACTTTGGCGCTGTACAAACGAGCGTGTGGCGCTTGACTCTG 780
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 781 GTTCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGCGGATTCGC 840
Qy 281 AlaAenTyrSerAenProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
Db 841 GCTAACTACTTAACCCACAGCACACGCGCTTCTGTTGTGGCACCATCTCTGAGCAAC 900
Qy 301 AspaAlaLeuArgAlaIleTrpGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320
Db 901 GATCGTTTACGTGGATTTGGGAACAAGAGCTGATGATATGCGCCAGCGTATTCAGCGT 960
Qy 321 MetArgGlnLeuPheValAenThrLeuGlnLysGlyAlaAenArgAspPheSerPhe 340
Db 961 ATGCGTCAGTTGTTTCGTCATACGCTGCAGGAAAGGCGCAACCGCGACTTCAGCTTT 1020
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360

Db 1021 ATCATCAACAGACGCGCATGTTCTCTCAGTGGCTGACAAAAGAACAGTCTGGCT 1080
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenValAlaGlyMet 380
Db 1081 CTGCGCGAAGAGTTTGGCGTATATCGGTTGCTTCTGTCGCGTAAATGTGCGCGGATG 1140
Qy 381 ThrProAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1141 ACACCAAGATAACATGGCTCGCTGTGCGAAGCGATTGTGCGAGTGTG 1188

RESULT 2

PCT-US05-47111-4
; Sequence 4, Application PC/TUS0547111
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: PCT/US05/47111
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 1241
; TYPE: DNA
; ORGANISM: Escherichia coli
PCT-US05-47111-4

Alignment Scores:
Pred. No.: 2,12e-237 Length: 1241
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 1

US-10-673-786A-2 (1-396) x PCT-US05-47111-4 (1-1241)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 25 ATGTTTGAGAACATTACCGCGCTCTGCGGACCCGATCTGGCGCTGGCGATCTGTTT 84
Qy 21 ArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValThrLysAspGluThr 40
Db 85 CGTCCGATGAACGTCCTGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 144
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 145 GGCACAAACCCGGTACTGACCGCGTGAAGGCTGAACAGTATCTGCTCGAAATGAA 204
Qy 61 ThrThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 205 ACCACCAAAATTTACCTCGGATTTGACGGCATCCCTGAATTTGCTGCTGCACTCAGGAA 264
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
Db 265 CTGCTGTTTGGTAAAGTAGCGCTGATCAATGACAAACGTCGTCGACGGCACAGACT 324
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 325 CCGGGGGGCACCTGGCGCACTACCGTGGCTGCCGATTTCTGGCAAAATTAACAGCGTT 384
Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 385 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGCAACCATAGAGCGCTTTTAACTCTGCA 444
Qy 141 GlyLeuGluValArgGluTyrAlaTyrArgAlaGlnAenHisThrLeuAspPheAsp 160
Db 445 GGTCTGGAAGTTCTGTAATACGCTTATATGATCGGGAATTAACACTCTTTGACTTCGAT 504
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 505 GCACTGATTAACAGCTGAAATGAAGCTCAGGCTGGCGAGCTAGTGTCTTCCATGGCTGC 564

QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 DB 565 TGCCATAAACCACCGATATGACCTACGCTGGAAACAATGCAACAACTGGCACAATC 624
 QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyGlnGlyPheAlaArgGly 220
 DB 625 TCCGTTGAGAAAGGCTGGTTACCGCTGTTGACTTCGCTTACACGGGTTTGGCCGTTGT 684
 QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 DB 685 CTGGAGAGAGATGCTGAAGGACTGGCGCTTTCGGGGCTATGCAATAGAGAGCTGATTGT 744
 QY 241 AlaSerSerTySerLysAsnPheGlyLeuTyAsnGluArgValGlyAlaCysThrLeu 260
 DB 745 GCCAGTTCTACTCTTAAACCTTTGGCCTGTACAAACGAGCGGTTTGGCGCTTGTACTCTG 804
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 DB 805 GTTGCTGGCGAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGCGGATTCCG 864
 QY 281 AlaAsnTySerAsnProAlaHisGlyAlaSerValAlaAlaThrIleLeuSerAsn 300
 DB 865 GCTAACTACTTAACCCACGACGACGCGGCTTCTGTTGTCACCACTCTGAGCAAC 924
 QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 DB 925 GATCGTTACGTCGATTGGGAAACAAGAGCTGATGATATGCGCCAGCGATTTCAGCGT 984
 QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
 DB 985 ATGCGTCAAGTTGTTGCTCAATACGCTGCAGGAAAAAGGCGCAACCCGCGACTTCA 1044
 QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 DB 1045 ATCATCAACAGACGAGCTATTCCTTCAGTGGCCCTGACAAAAAGAACAAAGTGTGCG 1104
 QY 361 LeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 DB 1105 CTGCGCGAAGAGTTGGCGTATATGCGGTGCTTCTGTCGCGTAAATGTGGCGGGATG 1164
 QY 391 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 DB 1165 ACACCAAGATATGCGCTGCGTGGAAAGGATTTGTGGCAGTCTG 1212

RESULT 3
 US-60-752-355-4
 ; Sequence 4, Application US/60752355
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark et al.
 ; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
 ; FILE REFERENCE: 38-21(53720)B
 ; CURRENT APPLICATION NUMBER: US/60/752,355
 ; CURRENT FILING DATE: 2005-12-21
 ; NUMBER OF SEQ ID NOS: 52803
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 4
 ; LENGTH: 1241
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 US-60-752-355-4

Alignment Scores:
 Pred. No.: 2,126-237 Length: 1241
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0
 US-10-673-786A-2 (1-396) x US-60-752-355-4 (1-1241)
 QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProfileLeuGlyLeuAlaAspLeuPhe 20
 |||

DB 25 ATGTTTGAGAACATTACCGCGCTCTCCGCAACCGATTCTGGGCGCTGGCCGATCTGTTT 84
 QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThr 40
 DB 85 CGTCCGATGAACGCTCCGCGCAAAATTAACTCTGGGATGGTGTCTATTAAGATGAGAGC 144
 QY 41 GlyIysThrProValLeuThrSerValLysLysAlaGluGlnTyLeuLeuGluAsnGlu 60
 DB 145 GGCAAAACCCCGGTACTGACGAGCGTGAAGGCGTGAACAGTATCTGCTCGAAATGAA 204
 QY 61 ThrThrLysAsnTyLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 DB 205 ACCACCAAAATTAACCTCGGCATTTGACGCGCATCCCTGAATTTGGTCGCTGCACTCAG 264
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 DB 265 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGCAAAACGCTGTCGACGCGCAGACT 324
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 DB 325 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCCGATTTCTGGCAAAAAATAACACGCGTT 384
 QY 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 DB 385 AAGCGTGTGTCGGTGAGCAACCCAGCTGGCGCAACCAATAAGAGCGTCTTTAACTCTGCA 444
 QY 141 GlyLeuGluValArgGluTyAlaTyTrpAspAlaGluAsnHisThrLeuAspPheAsp 160
 DB 445 GGTCTGGAAGTTCGTGAATACGCTTATGATGCGGAAATCAACACTCTTGACTTCGAT 504
 QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 DB 505 GCACGTGATTAACGCGTGAATGAAGCTCAGCGCTGCGAGCGTAGTCTGTTCCATGGCTGC 564
 QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 DB 565 TGCCATAAACCACCGGTATCGACCTACGCTGGAAACAATGCAACAACTGGCACAACATC 624
 QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyGlnGlyPheAlaArgGly 220
 DB 625 TCCGTTGAGAAAGGCTGGTTACCGCTGTTGACTTCGCTTACCGGGTTTCCCGCGTGT 684
 QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 DB 685 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCGCGGCTATGCAATAAGAGCTGATTGT 744
 QY 241 AlaSerSerTySerLysAsnPheGlyLeuTyAsnGluArgValGlyAlaCysThrLeu 260
 DB 745 GCCAGTTCTACTCTTAAACCTTTGGCCTGTACAAACGAGCGGTTTGGCGCTTGTACTCTG 804
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 DB 805 GTTGCTGGCGAGCTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGCGGATTCGC 864
 QY 281 AlaAsnTySerAsnProAlaHisGlyAlaSerValAlaAlaThrIleLeuSerAsn 300
 DB 865 GCTAACTACTTAACCCACGACGACGCGGCTTCTGTTGTCACCACTCTCTGAGCAAC 924
 QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 DB 925 GATCGTTACGTCGATTTGGGAAACAAGAGCTGATGATATGCGCCAGCGCTATTTCAGCGT 984
 QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
 DB 985 ATGCGTCAGTTGTTGCTCAATACCGTGGGAAAGGCGCAACCCGCGACTTTCAGCTTT 1044
 QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 DB 1045 ATCATCAACAGACGAGCTATTCCTTCAGTGGCCCTGACAAAAAGAACAAAGTGTGCG 1104
 QY 361 LeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 DB 1105 CTGCGCGAAGAGTTGGCGTATATGCGGTGCTTCTGTCGCGTAAATGTGGCGGGATG 1164

```
Qy 381 ThrProAspAenMetAlaProLeuGluCysGluAlaIleValAlaValLeu 396
Db 1165 ACACAGATAACATGGCTCCGCTGCGTGGCAAGCGATTGTGGCAGTGCTG 1212

RESULT 4
PCT-US06-00964-112
; Sequence 112, Application PC/TUS0600964
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: PCT/US06/00964
; CURRENT FILING DATE: 2006-01-23
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 112
; LENGTH: 1576
; TYPE: DNA
; ORGANISM: Escherichia coli K12
PCT-US06-00964-112

Alignment Scores:
Pred. No.: 1,63e-236 Length: 1576
Score: 2039.00 Matches: 395
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 1 Gaps: 0

US-10-673-786A-2 (1-396) x PCT-US06-00964-112 (1-1576)
Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 298 ATGATTGAGAACATTTACCGCGCTCTCCGACCGCATTCCTGGGCGCTGGCGATCTGTTT 357
Qy 21 ArgAlaAspGluAArgProGlyLysIleAenLeuGlyIleGlyValTyrLysAspGluThr 40
Db 358 CGTCCGATGAACGTCGCGCAAAATTAACCTCGGATTGGTGTCTATAAAGATGAGACG 417
Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 418 GCGAAACCCCGGTACTGACGAGCGTGAAAGGCTGAACAGTATCTCTCGAAATGGA 477
Qy 61 ThrThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyAArgCysThrGlnGlu 80
Db 478 ACCACAAAATTAACCTCGCATTCGCGCATCCCTGAATTTGGTGGTGGTGGTGGTGGT 537
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
Db 538 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCACGCGACAG 597
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 598 CCGGGGGGACCTGGCGCATACGCGTGGCTGCCGATTCCTGGCAAAATTAACCGCGTT 657
Qy 121 LysArgValTyrValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 658 AAGCGTGTGGTGAGCAACCCAGCTGGCGCAACCATAGAGCGTCTTTAACTCTGCA 717
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAspPheAsp 160
Db 718 GGTCTGGAAGTTCTGTGAATACGCTTATATGATGCGGAAATTCACACTCTTGACTTCG 777
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 778 GCACTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGATGCTGTTCATGGCTGC 837
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTyrTrpGlnThrLeuAlaGlnLeu 200
Db 838 TGCCATAACCCAAACCGGTATCGACCTACGCTGGAAACAATGGCAACACTGGCACAACTC 897
Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 941 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAenGlu 60

898 TCCGTTGAGAAAGCGCTGTTACCGCTCTTTGACTTCGCTTACCAGGTTTTTGGCCGCTGGT 957
221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
958 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCGCGCTATGCATAAAGAGCTGATTGTT 1017
241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyAlaCysThrLeu 260
1018 GCCAGTTCTTACTCTAAAAACTTTGGCCTGTGTACAAAGAGCGGTGTGGCGCTTGACTCTG 1077
261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
1078 GTTGTCTCCGACAGTGAACCGTTGATCGCGCATTCAGCAATGAAAGCGCGGATTCGC 1137
281 AlaAenTyrSerAenProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
1138 GCTAACTACTCTAACCCACCAGCACACGGCGCTTCTGTGTGTGGCACCATCTCTGAGCAAC 1197
301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
1198 GATGCGTTACGTGCGATTGGGAAACAAGAGCTGATGATGCGCCAGCGTATTACAGCGT 1257
321 MetArgGlnLeuPheValAenThrLeuGlnLysGlyAlaAenArgAspPheSerPhe 340
1258 ATGCGTCAGTTGTTCTGTCATATACGCTCAGGAAAGGCGCAACCGGACTTCAGCTTT 1317
341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
1318 ATCATCAACAGAACGCGCATGTTCTCTCTCAGTGGCTGACAAAAGAACAAAGTGTGCGT 1377
361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenValAlaGlyMet 380
1378 CTGCGCGAAGAGTTGGCGTATATGCGGTGCTTCTGTCGCTAAATGTGGCGGATG 1437
381 ThrProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
1438 ACACAGATAACATGGCTCCGCTGCTGCGAAGCGATTGTGGCAGTGCTG 1485

RESULT 5
US-11-330-403-112
; Sequence 112, Application US/11330403
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; CURRENT FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 112
; LENGTH: 1576
; TYPE: DNA
; ORGANISM: Escherichia coli K12
US-11-330-403-112

Alignment Scores:
Pred. No.: 1,63e-236 Length: 1576
Score: 2039.00 Matches: 395
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 8 Gaps: 0

US-10-673-786A-2 (1-396) x US-11-330-403-112 (1-1576)
Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 298 ATGATTGAGAACATTTACCGCGCTCTCCGACCGCATTCCTGGGCGCTGGCGATCTGTTT 357
Qy 21 ArgAlaAspGluAArgProGlyLysIleAenLeuGlyIleGlyValTyrLysAspGluThr 40
Db 358 CGTCCGATGAACGTCGCGCAAAATTAACCTCGGATTGGTGTCTATAAAGATGAGACG 417
Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 418 GCGAAACCCCGGTACTGACGAGCGTGAAAGGCTGAACAGTATCTCTCGAAATGGA 477
Qy 61 ThrThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyAArgCysThrGlnGlu 80
Db 478 ACCACAAAATTAACCTCGCATTCGCGCATCCCTGAATTTGGTGGTGGTGGTGGTGGT 537
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
Db 538 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCACGCGACAG 597
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 598 CCGGGGGGACCTGGCGCATACGCGTGGCTGCCGATTCCTGGCAAAATTAACCGCGTT 657
Qy 121 LysArgValTyrValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 658 AAGCGTGTGGTGAGCAACCCAGCTGGCGCAACCATAGAGCGTCTTTAACTCTGCA 717
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAspPheAsp 160
Db 718 GGTCTGGAAGTTCTGTGAATACGCTTATATGATGCGGAAATTCACACTCTTGACTTCG 777
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 778 GCACTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGATGCTGTTCATGGCTGC 837
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTyrTrpGlnThrLeuAlaGlnLeu 200
Db 838 TGCCATAACCCAAACCGGTATCGACCTACGCTGGAAACAATGGCAACACTGGCACAACTC 897
Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 941 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAenGlu 60
```



```
QY 142 LeuGluValArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAspPheAspAla 161
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 562552 ATCCAAGACAAACCTTATCGTTACTATGATGCCGCAACACGGTTTGGATTGGACGGC 562493

QY 162 LeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCys 181
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 562492 ATGATTAGGACCTTGAGCAAGCGCAAAAGCGCAATCGTCTCTCCACGGCTGTGTC 562433

QY 182 HisAenProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeuSer 201
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 562432 CACAATCTACCGGATGACCTACGCCGCAACCAATGGAAACCTTTGGCAAACTTTCT 562373

QY 202 ValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaAargGlyLeu 221
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 562372 GCCGAAAAGGCTGCTGCGCTGTTTGCATTTTGCCTACCAAGGCTTCGGCAATGTTTG 562313

QY 222 GluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleValAla 241
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 562312 GAAGAAGATGCCCTACGGCTTGGCGGTGTTTGTGAAACACAATACAGAAATGCTGATTGCC 562253

QY 242 SerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyValaCysThrLeuVal 261
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 562252 AGCTCTTATTCAAAACCTTCGTAAGTACAAACGAGCGCGTTCGCGCTTACATTGGTG 562193

QY 262 AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAla 281
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 562192 GCCGAAGATGAAGAAACAGCAGCCGCCGCCACACCAAGTCANAAACCATCATCGTACC 562133

QY 282 AenTyrSerAenProAlaHisGlyAlaSerValValAlaThrIleLeuSerAenAsp 301
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 562132 TTGTATTCCAAACCCGGCTTCACACGGTCGCAACACCATTTGGCTGCTTGAATAATGAT 562073

QY 302 AlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArgMet 321
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 562072 GATTTGAAAGCACAAATGGATTGCCGAATCGATGAAATGCGCGCGCGCATCAAGCCCATG 562013

QY 322 ArgGlnLeuPheValAenThrLeuGlnGlyLysGlyAlaAenArgAspPheSerPheIle 341
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 562012 CGCCAAAATTGTGGGTGCTCAAGCCCAAGGTGCAAGCCAAACCTTTGATTTCATT 561953

QY 342 IleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeu 361
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 561952 ATCAACAAACAAACGGTATGTTCTTTACGGCTTGAATCCCGAACAAGTCGACCGCGCTG 561893

QY 362 ArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenValAlaGlyMetThr 381
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 561892 AAAAAACGAGTTTGGCCATTATGCCCTCCGCTCCGCGCCGATCAACGTCGCGGCATTACC 561833

QY 382 ProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 561832 GACAACAACATCGATTATCTGTGCGAAGCATCGTGAAAGTACTG 561788
```

RESULT 7

```
PCT-US05-47111-133
; Sequence 133, Application PC/TUS0547111
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: PCT/US05/47111
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 133
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Zea mays
PCT-US05-47111-133
```

```
Alignment Scores: 5.37e-100 Length: 1230
Pred. No.: 913.00 Matches: 180
Score:
```

```
Percent Similarity: 62.6% Conservative: 69
Best Local Similarity: 45.2% Mismatches: 147
Query Match: 44.6% Indels: 2
DB: 1 Gaps: 2
```

US-10-673-786A-2 (1-396) x PCT-US05-47111-133 (1-1230)

```
QY 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 22 GTCTTCGCGCAGCTCGAGCAGGCCCGGAGACCCCATCTCTCGAGTGACCGTTGCTCTAC 81

QY 21 ArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValTyrLysAspGluThr 40
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 82 AACAAAGATCCAGCCCATGAAGGTCAACCTCGGGGTGCGGCTACCGGACCGAGGAA 141

QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 142 GGGAAAGCCCTAGTGTCTGAACGTGTGAGCGGCCGCGGCAAAATGTTGATCAATAATCCG 201

QY 61 Thr---ThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGln 79
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 202 TCACGTGTCAAGGAGTACCTACCAATCACCGGTCTGGCTGAATTCATAAAGCTGAGCGCT 261

QY 80 GluLeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGln 99
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 262 AAGCTTATCTTTGGCGCTGACAGCCCTGCTATTGAGGAGAAATAGGGTTGTACCGTGCGAG 321

QY 100 ThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSer 119
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 322 TGCCTATCGGGTACTGTTCTTTAAGAGTCGGAGTGCAATTTCTTGAAGGCACACTATCAC 381

QY 120 ValLysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSer 139
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 382 GAGCGCATCTACATCCACAAACCAACCTGGGGAAATCACCCAAAAGTCTTCACCCCTA 441

QY 140 AlaGlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAspPhe 159
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 442 TCTGGCTTGAACGTGTAGGAGTACCGCTATTATGATCTCTGCAACATGCGAGCTTCACTTC 501

QY 160 AspAlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGly 179
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 502 GAAGGACTCATGGGAAGACCTCGGTTCTCTTCAGGTTCAATTTGTTACTCTGCATGCC 561

QY 180 CysCysHisAenProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGln 199
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 562 TGTCCCAACACCCCTACTGAGTAGATCCTACCATCGAACACAGTGGGAACAGATTAGGCAG 621

QY 200 LeuSerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArg 219
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 622 CTGATGAGATCAAAATCACTGCTTCCGTTCTTTCACAGTGCCTATCAAGGCTTTGCAAGT 681

QY 220 Gly---LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeu 238
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 682 GGAAGCTTTGACAAAGATGTCAGTCAGTGGTATGTTTGTGCTGATGGTGGTGAACCTG 741

QY 239 IleValAlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyAlaCys 258
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 742 CTCATGCTCAGAGCTACGCTAAGAACATGGGATGTATGGAGAGCGGTGTTGGCGCTTTG 801

QY 259 ThrLeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAla 278
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 802 AGCATTTGTATAGAAAGTGGCGGATGTAGCTGTTAGGGTTGAAAGTCAACTCAAACTTGTG 861

QY 279 IleArgAlaAenTyrSerAenProAlaHisGlyAlaSerValAlaThrIleLeu 298
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 862 ATCAGGCTATGATTTCAAACCTCTCTTCATGCTGCTCTATTGTTGTGCTACCATCTC 921

QY 299 SerAenAspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIle 318
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 922 AGGACACGAGATGTTCAACGAATGGACTCTGGAACCTGGAAGGCGCATGGCTGATAGGATC 981

QY 319 GlnArgMetArgGlnLeuPheValAenThrLeuGlnGlyLysGlyAlaAenArgAspPhe 338
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
```

Db 982 ATTAACATGAGGCAACAACACTATTAAATGCGCTGAAATCCAGAGGAACCCCTGGTGATGG 1041
 Qy 339 SerPheIleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnVal 358
 Db 1042 AGCCATATCATTAAGCAATTTGGGATGTTTACTTTTCATGGCTGAATAGCAACAAGTG 1101
 Qy 359 LeuArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAla 378
 Db 1102 GCATTATCATGAGGAGCAATACACATTTATATGACATCTGATGGGAGGATCAGCATGGCC 1161
 Qy 379 GlyMetThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 1162 GGTITGAGCATGAGGAGTGTGCCCATCTTGCAGATGCCATACACGCTGCAGTT 1215
 RESULT 8
 US-60-752-355-133
 ; Sequence 133 Application US/60752355
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark et al.
 ; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
 ; FILE REFERENCE: 38-21(53720)B
 ; CURRENT APPLICATION NUMBER: US/60/752,355
 ; CURRENT FILING DATE: 2005-12-21
 ; NUMBER OF SEQ ID NOS: 52803
 ; SOFTWARE: Patent in version 3.3
 ; SEQ ID NO 133
 ; LENGTH: 1230
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; US-60-752-355-133
 Alignment Scores:
 Pred. No.: 5,378-100 Length: 1230
 Scores: 213.00 Matches: 180
 Percent Similarity: 62.6% Conservative: 69
 Best Local Similarity: 45.2% Mismatches: 147
 Query Match: 44.6% Indels: 2
 Gaps: 2
 DB: 12
 US-10-673-786A-2 (1-396) x US-60-752-355-133 (1-1230)
 Qy 1 MetPheGluAsnIleThrAlaAlaProAlaSerPheIleLeuGlyLeuAlaAspLeuPhe 20
 Db 22 GTCITTCGCGCACTCGAGCAGCCCGGAGGAGCCCATCTCGGAGTGACCGTTGCCCTAC 81
 Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 Db 82 AACAGGATCCAGCCCATGAAGTCAACCTCGGCGGTGGCGCTACCGGACCGAGGAA 141
 Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuAsnGlu 60
 Db 142 GGAAGCCCTAGTGTGAAGTGTGAGCGCGCGGAGCAATATGTTGATCAATATCCG 201
 Qy 61 Thr---ThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGln 79
 Db 202 TCACGTGTCAAGGATCTACCATCCCGTCTGGCTGAATTCATTAAGCTGAGCGCT 261
 Qy 80 GluLeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGln 99
 Db 262 AAGCTTATCTTGGCGCTGACAGCCCTGCTATTTCAGGAGATAGGTTGTCTACCGTGAG 321
 Qy 100 ThrProGlyGlyThrGlyValAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSer 119
 Db 322 TGCCTATCGGTACTGGTGTGTTAAGAGTGGAGGTGATTTCTTTCAGGAGCATATCAC 381
 Qy 120 ValLysArgValTyrValSerAsnProSerTyrProAsnHisLysSerValPheAsnSer 139
 Db 382 GAGCGCATCTATCATCCCAACACCACTGGGGAATCACCCAAAAGTCTTCCACCTA 441
 Qy 140 AlaGlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPhe 159
 Db 442 TCTGGCTTGAACGTTAGGAGCTACCGCTATTATGATCTCTGCAACATGACGCTTCACCTC 501

Qy 160 AspAlaIleLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGly 179
 Db 502 GAAGAGCTATGGAAGACCTCGGTTCTGCTCTCTCAAGTTCAATGTACTGTGTCATGCC 561
 Qy 180 CysCysHisAsnProThrGlyLysProThrGlyLeuGlnTyrGlnThrLeuAlaGln 199
 Db 562 TGTGCCCAACACCTTACTGGAGTAGATCTTACCATCGAACAGATGGGAACAGATTAGGCAG 621
 Qy 200 LeuSerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArg 219
 Db 622 CTGATGAGATCAAAATCACTGCTTCGCTTCTTGCAGCTGCTATCAAGGCTTTCGAGT 681
 Qy 220 Gly---LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeu 238
 Db 682 GGAAGTCTTGACAAGATGCTCAGTCAGTGGCTATGTTTGTGCTGATGCTGGTGAACATG 741
 Qy 239 IleValAlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGluValaCys 258
 Db 742 CTCATGGCTCAGAGCTACGCTAAGAACATGGATTTGATGGAGAGCGTGTGGCCCTTGG 801
 Qy 259 ThrLeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAla 278
 Db 802 AGCATTTGATGATAGAGTCCGATGCTAGCTTTAGGGTTGAAGTCAACTCAACTTGTCTC 861
 Qy 279 IleArgAlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeu 298
 Db 862 ATCAGGCTATGATTTCAACCTCTCTCATGCTGCTCTATTGTTGCTACCATCTC 921
 Qy 299 SerAsnAspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIle 318
 Db 922 AGGACAGCAGAGATGTTCAACGAATGGACTCTGGAACCTGAAGCCATCGCTGATAGGATC 981
 Qy 319 GlnArgMetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPhe 338
 Db 982 ATTAACATGAGCAACAACACTATTAAATCGCTGAAATCCAGAGGAACCCCTGGTATGG 1041
 Qy 339 SerPheIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnVal 358
 Db 1042 AGCCATATCATTAAGCAATTTGGGATGTTTACTTTCTGCTGCTGAATAGCAACAAGTG 1101
 Qy 359 LeuArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAla 378
 Db 1102 GCATTATCATGAGGAGCAATACACATTTATATGATCATCTGATGGGAGGATCAGCATGCC 1161
 Qy 379 GlyMetThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 1162 GGTITGAGCATGAGGAGTGTGCCCATCTTGCAGATGCCATACACGCTGCAGTT 1215
 RESULT 9
 PCT-US05-47111-446
 ; Sequence 446 Application PC/TUS0547111
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark et al.
 ; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
 ; FILE REFERENCE: 38-21(53720)B
 ; CURRENT APPLICATION NUMBER: PCT/US05/47111
 ; CURRENT FILING DATE: 2005-12-21
 ; NUMBER OF SEQ ID NOS: 52803
 ; SOFTWARE: Patent in version 3.3
 ; SEQ ID NO 446
 ; LENGTH: 1224
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; PCT-US05-47111-446
 Alignment Scores:
 Pred. No.: 1,638-99 Length: 1224
 Scores: 909.00 Matches: 179
 Percent Similarity: 62.3% Conservative: 69
 Best Local Similarity: 45.0% Mismatches: 148
 Query Match: 44.4% Indels: 2
 Gaps: 2
 DB: 1


```

Alignment Scores:
Pred. NO.:          2.13e-96      Length:      1668
Score:              885.00        Matches:    180
Percent Similarity: 61.1%         Conservative: 66
Best Local Similarity: 45.1%       Mismatches: 149
Query Match:        43.3%         Indels:     4
DB:                  7            Gaps:       3

US-10-673-786A-2 (1-396) x US-10-953-349-16412 (1-1668)

Qy      1 MetPheGluAgnIleThrAlaAlaProAlaaSPProfileLeuglyLeualaaspLeupHe 20
      :|::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
      :|::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

314	Db	GTCTTCAATCACTCTGGTTCTGCTCCGGAAGATCCTATCTCTCGGGGTAACTGTGCGCTTAT	373
21	Qy	ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr	40
374	Db	AACAAAGATCCAAGTCCAGTTAAGCTCAACTTGGAGTTGGTCTTACCGAACTGAGAA	433
41	Qy	GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu	60
434	Db	GGAAAACTCTGTGTTTGAATGTAGTGAGCGGAGTTGAACAGCAACTCATATA--AATGAC	490
61	Qy	ThrThrLysAsn-----TyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThr	78
491	Db	GTGTGCGCAACCAAGGAATATATCCGATCGTTGGGCTTCTGCTGAITTAATAAATCGAT	550
79	Qy	GlnGluLeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAla	98
551	Db	GCTAAGCTTATTTTGGGCTGACAGCCCTGCTATTCAACACAACAGGGTTACCACCTGT	610
99	Qy	GlnThrProGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThr	118
611	Db	CAATGCTTGTCTGGAATCGTCTTTAAGAGTTGGGGTGAAATTTTGGCTTAAACACTAT	670
119	Qy	SerValLysArgValTyrPValSerAsnProSerTyrProAsnHisLysSerValPheAsn	138
671	Db	CACCAACGGCATATATATCTGCCAACACCACTTGGGGCAATCACCAGAGGTTTTCACAC	730
139	Qy	SerAlaGlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAsp	158
731	Db	TTAGCAGCCTTGTCTCAAAAACATACCGCTACTATGCTCCAGCAACACAGGACTTGAC	790
159	Qy	PheAspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHis	178
791	Db	TTTCAAGGACTTCTGGAAGACCTTGGTTCTGCTCCATCTGATCTATTGTTTGTGTACAT	850
179	Qy	GlyCysCysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAla	198
851	Db	GCAATGGGCATATAACCCACTGTGTGGATGCCAACCCCTTGACGAATGGGAGAGATTAGG	910
199	Qy	GlnLeuSerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAla	218
911	Db	CAGCTAATAGATCAAAAGCTTGTGTACCTTTCTTTGACAGTCTTATCAGGGTTTGTCT	970
219	Qy	ArgGly----LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGlu	237
971	Db	AGTGGAGTCTAGATGCAGATGCCAACCTGTTCTGTTTGTGTGATGAGGCGCAA	1030
238	Qy	LeuIleValAlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAla	257
1031	Db	TTGCTGTGTACCAAAAGCTATGCAAGAATCTGGGTCTTTATGGGGAACGTGTGGCGCC	1090
258	Qy	CysThrLeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAla	277
1091	Db	TTAAGCATTTCTGCAAGTCAGCTGATGTTGCAAGCAGGGTTGAGAGCCAGCTGAAGCTA	1150
278	Qy	AlaIleArgAlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIle	297
1151	Db	GTGATTAGGCCCATGTACTCAAGTCTCCCATTCATGTTGTCATCCATTTGGCTGCCATT	1210
298	Qy	LeuSerAsnAspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArg	317
1211	Db	CTCAAGACCGGAATTTGTTTCAATGACTGGACTATTGAGTTGAAGGCAATGGCTGATCGC	1270
318	Qy	IleGlnArgMetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAsp	337
1271	Db	ATCATCGATGATGCCCAAGAACTTTTCGATGCTTTATGTTCCAGAGCGCACACCTGGCGAT	1330
338	Qy	PheSerPheIleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGln	357
1331	Db	TGGAGTCACATTATCAACACAGATTGGAAATGTTTACTTCTTCACTGGATTGAATCGGAACAA	1390
358	Qy	ValLeuArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnVal	377
1391	Db	GTTTCTTCACTAGTCAAAAGATTTCCATATATACATGACATCTCATGGGAGATTAGCATG	1450

Percent Similarity:	59.0%	Conservative:	73
Best Local Similarity:	40.7%	Mismatches:	161
Query Match:	39.8%	Indels:	2
DB:	10	Gaps:	2

US-10-673-786A-2 (1-396) x US-11-299-286-563 (1-1218)

1 MetPheGluAsnIleThrAlaIaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db ATTCCTGCAAGCGTCTCTCTCTCCCAAGATCTCTGTTCAGTGAATATTTTGTCTGT 69
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThr 40
Db AGAGATGATCCTAGTCCGGTTAAGTTGAATTTGACGCGAGGTACCTATCGAACTGAGGA 129
Qy 41 GlyTyThrProValLeuThrSerValLysIleAlaGluInTyLysLeuGluAsn--- 59
Db GGAAAGCCCTCTTGTCTTGATGTTGTGAAGAGCAGACCAACATGTATGACCAACGACCTG 189
Qy 60 GluThrThrLysAsnTyLysLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGln 79
Db TCTAGGACAGGAATATCTCCCTCAATGGACTTCTGTGATTTAATAATTAAGTACC 249
Qy 80 GluLeuLeuPheGlyLysGlySerAlaLeuIleAsnAspTyLysArgAlaArgThrAlaGln 99
Db AAGCTCATCTTAGTTCATGATAGTCTCTGCATTGAAAGAGATAGAGTTGTTTACAAACCAA 309
Qy 100 ThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSer 119
Db TGTTCCTGCTGGTACTGGTTCCTTTGAGAGTTGGAGCTGAGTTTCAGCAACACACAA 369
Qy 120 ValLysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSer 139
Db GAAAGTGTCATTTTCGTTCCCAACCAACGTTGGGGGACCATCCCGCTATTTTACTTTG 429
Qy 140 AlaGlyLeuGluValArgGluTyLysAlaTyTrpAspAlaGluAsnHisThrLeuAspPhe 159
Db GCGGGTTCTCTGTCCAGTATTTTCGTTATTTATGATCGAAGCCGAGGACTCGACTTC 489
Qy 160 AspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGly 179
Db AAGGCATGCTCGAGATCTTGGCGCTGCACCTCTCGAGCTATAGTGGTACTTCAAGCT 549
Qy 180 CysCysHisAsnProThrGlyIleAspProThrLeuGluGlnInTyLysLeuAlaGln 199
Db TGTGCGCATAAACCCACAGGATTCACCAACATTCGAAACAATGGGAAAGATTCGACGA 609
Qy 200 LeuSerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyTrpGlnGlyPheAlaArg 219
Db CTAGTGAGATCGAAAGCTTATACCTTCTTTGATGTGCATATCAGGGTTTGTCTAGT 659
Qy 220 Gly---LeuGluGluAspAlaGlyLeuArgAlaPheAlaMetHisLysGluLeu 238
Db GGTAGCCTTGACGCAGATGCACAAGCGGTTCTGTATGTTGTAGCCGATGGAGGTGAATG 729
Qy 239 IleValAlaSerSerTyLysAsnPheGlyLeuTyTrpAsnGluArgValGlyAlaCys 258
Db TTGATAGCTCAAAAGTTATGCCAAAATATGGGCTTTATGGGAGCGATTTGGCTCTCTT 789
Qy 259 ThrLeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAla 278
Db ACGATCGTATGCATCATCAGATGTCGTAAGAAAGTTGAGACCAAGTCTACTTGT 849
Qy 279 IleArgAlaAsnTyLysSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeu 298
Db GTGAGCGCTATGTATCTTACACACCTATTTCATGTGTGCATCTATTGTTGGCCAAATCTA 909
Qy 299 SerAsnAspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIle 318
Db AAAAAAGTATGATGATCAATGACTGGACCATTTAGCTTAAAGGANTGGCTGACCGGATA 969
Qy 319 GlnArgMetArgGlnLeuPheValAsnThrLeuGlnGlyLysGlyAlaAsnArgAspPhe 338

970	ATTAGCATGGCTCAACAGTTATATGACGCTCTAGAGCTAGAGCAACACCTGGTGATTGG	Db
339	SerPheIleIleIleYsGlnAsnGlyMetPheSerPheSerGlyLeuThrIysGluGlnVal	Qy
1030	AGTCACATTATTAACACATATTGGATGTTTACTTTTACCGGATTGATGAGGAGCAAGTT	Db
359	LeuArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAla	Qy
1090	CGCTTGATGCCAAAGAATATCATTTTACATGACTTTATGATGGGAGGATAAGCATGGCA	Db
379	GlyMetThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu	Qy
1150	AGTCTAAGTTCGAAGACAGTCGCTCAACTCGCTGATGCTATACATGCTGTGTGT	Db

RESULT 14

```

US-10-953-349-11025
; Sequence 11025, Application US/10953349
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED D
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 11025
; LENGTH: 1301
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-11025

```

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

US-10-673-786A-2 (1-396) x US-10-953-349-11025 (1-1301)

Qy	1	MetPheGluAenIleThrAlaAalaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe	20
Db	17	ATCTTGCAGGGCTCTCTCGCTCCCGAAGATCCTGTCTCAGTGAATATTTTGCCTGT	76
Qy	21	ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThr	40
Db	77	AGAGATGATCTAGTCGGTTAAGTTGAATTTGACCGCAGGTACCTATCGAACTGAGGA	136
Qy	41	GlyLysThrProValLeuThrSerValbysIysAlaGluGlnTyLeuLeuGluAen	59
Db	137	GGAAAGCCCTCTGTCTTGATGTTGTGAGAAAGACAGACACAGTTAGCCCAACGACCTG	196
Qy	60	GluThrThrLysAenTyLeuGlyIleProGluPheGlyArgCysThrGln	79
Db	197	TCTAGGAGACAGGAATATCTCCCTCAATGACATTCCTGAAATTTAATAAATTAAGTACC	256
Qy	80	GluLeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGln	99
Db	257	AAGCTCATCTTAGGTGATGATGCTCTCATTTGAAAGAGAATAGAGTTGTTTACAACCCAA	316
Qy	100	ThrProGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSer	119
Db	317	TGTTTGCTGTGCTACTGGTCTTTGAGAGTTGAGCTGAGTTTCTAGACACACACAACAA	376
Qy	120	ValLysArgValTrpValSerAsnProSerTrpProAenHisLysSerValPheAenSer	139
Db	377	GAATGTGTCATTTTCGTTCCAAACCACTTGGGGGAACCATCCCGTATTTTACTTTTG	436
Qy	140	AlaGlyLeuGluValArgGluTyAlaTyTrpAspAlaGluAAsnHisThrLeuAspPhe	159
Db	437	GCGGTTTTGCTGTCGAGTATTTCCGTTACTATGATCCGAAGGCGGACACTCGATTTTC	496

```
Qy 160 AspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGly 179
Db : : : : :
497 AAAGCATGCTCGAGGATCTTTGGCGCTCGACCTCTCTGGAGCTATAGTGGTACTTCAAGCT 556
Qy 180 CysCysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGln 199
Db : : : : :
557 TGTGGCATAACCCACAGGAGTTGACCCACAATCGCACAAATGGGAACACGATTGCGAGA 616
Qy 200 LeuSerValGluLysGlyTrpLeuProLeuPhePheAlaTrpGlnGlyPheAlaArg 219
Db : : : : :
617 CTAGTGACATCGAAGAGCTTATTACCTCTTTGATAGTGATATACAGGTTTGTCTAGT 676
Qy 220 Gly---LeuGluGluAspAlaGlyLeuArgAlaPheAlaAlaMetHisLysGluLeu 238
Db : : : : :
677 GGTAGCCTTGACGAGATGACCAAGCGGTCGTATGTTGTCGATGGAGGTGAATGT 736
Qy 239 IleValAlaSerSerTySerLysAsnPheGlyLeuTyAsnGluArgValGlyAlaCys 258
Db : : : : :
737 TTGATAGCTCAAGTTATGCCAAAATATGGCCCTTTATGGCGAGCGTATTGGCTCTCT 796
Qy 259 ThrLeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAla 278
Db : : : : :
797 AGATCGTATGACACATCAGAGATGCTCGTAAAGAGTTGAGAACCAAGTCTACTGTGT 856
Qy 279 IleArgAlaAsnTySerAsnProAlaHisGlyAlaSerValValAlaThrIleLeu 298
Db : : : : :
857 GTGAGGCTATGATCTTACACCACTATTATGCTGTCATCTATTGTTGCCACAATCTA 916
Qy 299 SerAsnAspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIle 318
Db : : : : :
917 AAAAAACAGTATATGTACAATGACTGGACCACTTGGCTGAAAGGAATGGCTGACGCGCATA 976
Qy 319 GlnArgMetArgGlnLeuPheValAsnThrLeuGlnLysGlyAlaAsnArgAspPhe 338
Db : : : : :
977 ATTAGTAGCTGCGTCAACAGCTTATATGACGCTCTAGAAGCTAGAGGAAACACCTGGTGATGG 1036
Qy 339 SerPheIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnVal 358
Db : : : : :
1037 AGTCACATTATTAACATATTGGGATGTTTACTTTTACCGGATTGAGTGAGGACGAGTT 1096
Qy 359 LeuArgLeuArgGluLupPheGlyValTyTrpAlaValAlaSerGlyArgValAsnValAla 378
Db : : : : :
1097 CGCTTGATGGCCAAAGAATATACATTTTACATGACTATGATGGGAGGATAAGCATGGCA 1156
Qy 379 GlyMetThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db : : : : :
1157 AGTCTAAGTTCGACGACAGTGCCTCAACTCGCTGATGCTATACATGCTGTGTGT 1210
```

RESULT 15

```
US-11-329-121-902
; Sequence 902, Application US/11329121
; GENERAL INFORMATION:
; APPLICANT: BIRSE, Charlie et al.
; TITLE OF INVENTION: ADIPOSE DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001565CDR
; CURRENT APPLICATION NUMBER: US/11/329,121
; CURRENT FILING DATE: 2006-01-11
; NUMBER OF SEQ ID NOS: 1724
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 902
; LENGTH: 2339
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-329-121-902
```

```
Alignment Scores:
Pred. No.: 3,95e-85 Length: 2339
Score: 794.00 Matches: 161
Percent Similarity: 58.3% Conservative: 67
Best Local Similarity: 41.2% Mismatches: 161
Query Match: 38.8% Indels: 2
DB: 8 Gaps: 2
```

```
US-10-673-786A-2 (1-396) x US-11-329-121-902 (1-2339)
Qy 4 AsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPheArgAlaAsp 23
Db : : : : :
111 CATGTGGAATGGGACCTCCAGATCCCATCTCTGGAGTCACTGAAGCCTTTAAGAGGAG 170
Qy 24 GluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThrGlyLysThr 43
Db : : : : :
171 ACCAATAGCAAAAAGATGAATCTGGGAGTTGGTGCCTACCGGATGATATGGAAGCCT 230
Qy 44 ProValLeuThrSerValLysLysAlaGluGlnTyTrpLeuLeuGluAsnGluThrThrLys 63
Db : : : : :
231 TACGTTCTGCTAGCTGCCAAGGAGAGGCCAGATTCGCCGAAAAAATTTGGACAAAG 290
Qy 64 AsnTyTrpLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGluLeuPhe 83
Db : : : : :
291 GAATACCTGCCATTGGGGGACTGGCTGAATTTTGCAGGACATCTGCAGAACTAGCCCTG 350
Qy 84 GlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThrProGlyGly 103
Db : : : : :
351 GGTGAGAACAGCGAAGTCTTGAAGAGTGGCGGTTTGTCACTGTGCAGACCATTTCTGGA 410
Qy 104 ThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerValLysArg--- 122
Db : : : : :
411 ACTGGAGCCTTAAAGGATCGGAGCCAGTCTTCTGCAAGATTTTTTAAGTTTCAAGCGAGAT 470
Qy 123 ValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAlaGlyLeu 142
Db : : : : :
471 GTCTTCTGCCAAACCACTGGGAAACCAACACACCCATCTTCAGGGATGCTGGCATG 530
Qy 143 GluValArgGluTyAlaTyTrpAspAlaGluAsnHisThrLeuAspPheAspAlaLeu 162
Db : : : : :
531 CAGCTACAAGGTTATCGTATTATGACCCCAAGACTTGGCGTTTGTGACTTCACAGGCGCT 590
Qy 163 IleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysHis 182
Db : : : : :
591 GTGAGGATATTTCAAAAATACCAGAGCAGAGTGTCTTCTTCTGTCATGCTGCGCCAC 650
Qy 183 AsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeuSerVal 202
Db : : : : :
651 AATCCCAAGGAGTGGACCCGCTCGGAACAGTGGGAAGGAATAGCACACAGTGGTGAAG 710
Qy 203 GluLysGlyTrpLeuProLeuPheAspPheAlaTyTrpGlnGlyPheAlaArgLys---Leu 221
Db : : : : :
711 AAAAGGAATCTCTTGGTGTCTTTGACATGCGCTTACCAAGCTTTGCGAGTGTGATGCT 770
Qy 222 GluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleValAla 241
Db : : : : :
771 GATAAGGATGCTGGGCTGTGCGCCACTTCATCGAACAGGGCATTAATGTTTGCCTTCGC 830
Qy 242 SerSerTyTrpSerLysAsnPheGlyLeuTyAsnGluArgValGlyAlaCysThrLeuVal 261
Db : : : : :
831 CAATCATATGCCAAGAACATGGGCTTATATGTTGGTGGCGTGTAGGAGGCTTCACCTATGCTC 890
Qy 262 AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAla 281
Db : : : : :
891 TGCAAGATGCGGATGAAGCCAAAAGGTAGAGTCACAGTTGAAGATCTTGATCCGCTCCC 950
Qy 282 AsnTyTrpSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsnAsp 301
Db : : : : :
951 ATGTATTCCAACCTCCCTCAATGGGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1010
Qy 302 AlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArgMet 321
Db : : : : :
1011 GATTTCGAAAACAAATGGCTGCAAGAGTGAAGGCGATGGCTGACCGGATCATTTGGCATG 1070
Qy 322 ArgGlnLeuPheValAsnThrLeuGlnGlyLysAlaAsnArgAspPheSerPheIle 341
Db : : : : :
1071 CGGACTCACTGGTCTCCAACCTCAAGAGGAGGTTCCACCCACAAATTTGGCAACATC 1130
Qy 342 IleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeu 361
Db : : : : :
```

Db 1131 ACCGACCAATGGCATGTTCTGTTTCACAGGGCTAAGCCTGAACAGGTGGAGCGGCTG 1190
 Qy 362 ArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
 Db 1191 ATCAAGGAGTTCTCCATCTACATGACAAAAGATGGCCGCATCTCTGTGGCAGGGGTCACC 1250
 Qy 382 ProAspAsnMetAlaProLeuCysGluAlaIle 392
 Db 1251 TCCAGCAACGTGGGCTACCTTGCCCATGCCATT 1283

Search completed: March 15, 2006, 16:06:44
 Job time : 1265 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM nucleic - nucleic search, using sw model
Run on: March 15, 2006, 14:30:49 ; Search time 1436 Seconds
(without alignments)
1933.887 Million cell updates/sec
Title: US-10-673-786A-1
Perfect score: 1191
Sequence: 1 atgttgagaacattaccgc.....cgattggcagctgtgtaa 1191
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 8023312 seqs, 1165852854 residues
Total number of hits satisfying chosen parameters: 16046624
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA_New:*
1: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
5: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
7: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
9: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
11: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
12: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
13: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1176.6	98.8	1191	12	US-11-114-922-75 Sequence 75, Appl
2	461.4	38.7	1269	8	US-10-467-657-6535 Sequence 6535, Ap
3	285.4	24.0	1191	8	US-10-467-657-1059 Sequence 1059, Ap
4	255.4	21.4	1194	8	US-10-979-821-31 Sequence 31, Appl
5	255.4	21.4	1194	12	US-11-114-922-31 Sequence 31, Appl
6	241.6	20.3	1260	8	US-10-979-821-5 Sequence 5, Appli
7	241.6	20.3	1260	12	US-11-114-922-5 Sequence 5, Appli
8	240	20.2	1260	8	US-10-979-821-3 Sequence 3, Appli
9	240	20.2	1260	12	US-11-114-922-3 Sequence 3, Appli
10	221.6	18.6	2343	12	US-11-136-527-2278 Sequence 2278, Ap
11	206.2	17.3	1170	8	US-10-979-821-1 Sequence 1, Appli
12	206.2	17.3	1170	12	US-11-114-922-1 Sequence 1, Appli
13	145.8	12.2	1239	8	US-10-979-821-7 Sequence 7, Appli
14	145.8	12.2	1239	12	US-11-114-922-7 Sequence 7, Appli
15	145.2	12.2	1251	7	US-10-932-182A-82350 Sequence 82350, A
16	145.2	12.2	1251	7	US-10-932-182A-82350 Sequence 82350, A
17	127.6	10.7	1389	7	US-10-932-182A-4671 Sequence 4671, Ap
18	127.6	10.7	1389	7	US-10-932-182A-4671 Sequence 4671, Ap
19	102	8.6	1400	12	US-11-136-527-6374 Sequence 6374, Ap
20	85	7.1	560	6	US-09-925-065A-479173 Sequence 479173,

ALIGNMENTS

RESULT 1

US-11-114-922-75
; Sequence 75, Application US/11114922
; Publication No. US20050282260A1
; GENERAL INFORMATION:
; APPLICANT: MCFARLAN, SARA C.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; PRIOR FILING DATE: 2005-04-26
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 75
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
US-11-114-922-75

Query Match	98.8%;	Score	1176.6;	DB	12;	Length	1191;
Best Local Similarity	99.2%;	Pred. No.	0;				
Matches	1182;	Conservative	0;	Mismatches	9;	Indels	0;
Gaps	0;						
QY	1	ATGTTTGAAGAACATTACCGCCGCTCTCGCGACCCCGAATTCGGGCCCTGGCCGATCTGTTTT	60				
DB	1	ATGTTTGAAGAACATTACCGCCGCTCTCGCGACCCCGAATTCGGGCCCTGGCCGATCTGTTTT	60				
QY	61	CGTGGCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTGTCTTATTAAGATGAGACG	120				
DB	61	CGTGGCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTGTCTTATTAAGATGAGACG	120				
QY	121	GGCAAAACCCCGGTACTGACACGCGTGAAGAAAGGCTGAACAGTATCTGCTCGAAATGAA	180				
DB	121	GGCAAAATCCCGGTACTGACACGCGTGAAGAAAGGCTGAACAGTATCTGCTCGAAATGAA	180				
QY	181	ACCACCAAAATTAACCTCGGCAATTAACGCAATTCCTGAAATTTGGTTCGCTCACTCAGGAA	240				
DB	181	ACCACCAAACTTACCTCGGCAATTAACGCAATTCCTGAAATTTGGTTCGCTCACTCAGGAA	240				


```

QY 241 CTGCTGTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTGCGCGCACAGACT 300
DB 241 CTGCTGTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTGCGCGCACAGACT 300

QY 301 CCGGGGGCACTGGCGCACTACGCGTGGCTGCCGATTTCTTGGCAAAAAATACACGGTT 360
DB 301 CCGGGGGCTCTGGCGCACTACGCGTGGCTGCCGATTTCTTGGCAAAAAATACACGGTT 360

QY 361 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGAAACCAATAGAGCGTCTTTAACTGTGCA 420
DB 361 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGAAACCAATAGAGCGTCTTTAACTGTGCA 420

QY 421 GGTCTGCAAGTCTGTGATAGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
DB 421 GGTCTGCAAGTCTGTGATAGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480

QY 481 GCACGTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACCTAGTGTCTTCCATGGCTGC 540
DB 481 GCACGTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACCTAGTGTCTTCCATGGCTGC 540

QY 541 TGGCATTAACCCAAACCGGTATCGACCCCTACGCTGGAAACATGGCAACACATGGCAACATC 600
DB 541 TGGCATTAACCCAAACCGGTATCGACCCCTACGCTGGAAACATGGCAACACATGGCAACATC 600

QY 601 TCGGTTGAGAAAGGCTGGTTACCGCTGTTTGAATTCGCTTACCAGGGTTTTGGCCGTGT 660
DB 601 TCGGTTGAGAAAGGCTGGTTACCGCTGTTTGAATTCGCTTACCAGGGTTTTGGCCGTGT 660

QY 661 CTGGAAGAAGTGTGAAAGACTGCGCGCTTTTGGCGCTATGCAATAAGAGCTGATGTT 720
DB 661 CTGGAAGAAGTGTGAAAGACTGCGCGCTTTTGGCGCTATGCAATAAGAGCTGATGTT 720

QY 721 GCCAGTTCTTACTCTTAAAACTTTGGCTGTGACAAAGAGCGTGTGGCGCTTGTACTCTG 780
DB 721 GCCAGTTCTTACTCTTAAAACTTTGGCTGTGACAAAGAGCGTGTGGCGCTTGTACTCTG 780

QY 781 GTTGCTGCGACAGTGAACCGTTGATCGCGCATTCAGCAAAATGAAGCGCGATTCGC 840
DB 781 GTTGCTGCGACAGTGAACCGTTGATCGCGCATTCAGCAAAATGAAGCGCGATTCGC 840

QY 841 GCTAACTACTCTTAACCCACAGCACAGCGCTTCTGTGTGCGCAATTCAGCAAAATGAAGCGCGATTCGC 900
DB 841 GCTAACTACTCTTCCCAACAGCACAGCGCGCTTCTGTGTGCGCAATTCAGCAAAATGAAGCGCGATTCGC 900

QY 901 GATGCGTTACGTGCGATTTGGGAAACAAGAGCTGATGATATGCGCCAGCGTATTCAGCGT 960
DB 901 GATGCGTTACGTGCGATTTGGGAAACAAGAGCTGATGATATGCGCCAGCGTATTCAGCGT 960

QY 961 ATGCGTCAGTTGTTCTGTCATACGCTGAGGAAAGCGGCAACCGGACTTCAGCTTT 1020
DB 961 ATGCGTCAGTTGTTCTGTCATACGCTGAGGAAAGCGGCAACCGGACTTCAGCTTT 1020

QY 1021 ATCATCAACAGAAAGCGCATGTTCTCTTCTGAGTGGCTGACAAAGAAACAAGTGTGCGT 1080
DB 1021 ATCATCAACAGAAAGCGCATGTTCTCTTCTGAGTGGCTGACAAAGAAACAAGTGTGCGT 1080

QY 1081 CTGCGGAAAGATTTGCGGTATATGCGGTTGCTTCTGCTGCGGTAATGTTGGCGGGATG 1140
DB 1081 CTGCGGAAAGATTTGCGGTATATGCGGTTGCTTCTGCTGCGGTAATGTTGGCGGGATG 1140

QY 1141 ACACGATTAACATGGCTCGCTGTGGAAGCGATTTGTGGCAGTGTGTTAA 1191
DB 1141 ACACGATTAACATGGCTCGCTGTGGAAGCGATTTGTGGCAGTGTGTTAA 1191

```

RESULT 2

```

; Sequence 6535, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita

```

```

; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6535
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; US-10-467-657-6535

```

```

Query Match 38.7%; Score 461.4; DB 8; Length 1269;
Best Local Similarity 61.9%; Pred. No. 3,7e-140;
Matches 732; Conservative 0; Mismatches 451; Indels 0; Gaps 0;

```

```

QY 2 TGTTTGAGAACATTACCGCGCTCTCGCGACCCGATTTCTGGCCTGGCGATCTGTTTC 61
DB 86 TCTTCAAGCACATCGAAGCGCCGCCGATCCGATTTCTCGGTTTGGGCGAAGCGTTCA 145

QY 62 GTGCCGATGAAGTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACGG 121
DB 146 AGCCGAAACCCGCCGCAAAAGTCAACCTCGGATCGGCTGTACAAAGACGATCCG 205

QY 122 GCAAAACCCCGGTACTGACGCGTGAAGAGGCTGAACAGTATCTGTCGAAAAATGAAA 181
DB 206 GCGCGACACCCATTGTCAAAGCGCTCAAAGAGGCTGAAAAACGCTGTTGGAAGCGAAA 265

QY 182 CACCAAAATTAACCTCGGATTTGACGGCATCCCTGAATTTGTCGCTGCATCAGGAAC 241
DB 266 CCACCAAAATTAACCTGCATCGACGCGTTTGCAGCTACAAAGAGCAACCCCAATTC 325

QY 242 TGTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCAGCGCACAGACTC 301
DB 326 TGTGTTTGGCAAGAACACGAAATCATGCCAGCGCGCGCAAAACAGCGCAAGGCC 385

QY 302 CGGGGGCACTGGCGCACTACGCGTGGCTGCCGATTTCTTGGCAAAAAATACAGCGTTA 361
DB 386 TTGGCGGTACAGCGCATTTGCGGTATTGCGGCGGAAATTTGCAAAACGTCAGTTGAAACGCGC 445

QY 362 AGCGTGTGGTGAACCAACCGTACCGGCGGCAACCATTAAGAGCGTCTTTAACTCTGCAG 421
DB 446 AAACCATCTGGATTTTCCAATCCGACCTGGCCAAACCAACCGCAATTTGCCAAAGCGGTGC 505

QY 422 GTCTGAAGTTCGTGAATACGCTTATATGATGCGGAAATACACACTCTTTGACTTCGATG 481
DB 506 GTATCCAGACAAACCTTATCGCTACTATGATGCCGCCAAACACTGCTTGGATTGGGACG 565

QY 482 CACTGATTAACAGCTGAAATGAAGCTCAGGCTGGCGACGTPAGTGTGTTCCATGGGTGCT 541
DB 566 GCATGATTGAAGTTTGAACCAAGCGCAAAAGAGCGACATCGTCTCTGCACGGGTGCT 625

QY 542 GCATTAACCAACCGGTATGACCCCTACGCTGGAAACATGGCAAACTGGCACTGCTCT 601
DB 626 GCCAACACCCCTACCGGTATCGACCTACGCGGCAACATGGGAAACTTTTAGCAAAACTTT 685

QY 602 CCGTTGAGAAAGGCTGTTACCGCTGTTTGAATTCGCTTACCGAGGTTTGGCCGTTGCT 661
DB 686 CTCCGAAAGAGGCTGTTGCGCATTTGTTGACTTTGCTTACCAAGGCTTCGGCAATGCTT 745

QY 662 TGGAAAGAGTGTGAAGGATGCGCGCTTTTCCGCGCTATGCGGCTATGCAATAAGAGCTGAT 721
DB 746 TGGAAAGAGTGTGCTATGGCTTTCGCGCTGTTCTTTGAAACACAATACAGAATTCGTGATTG 805

QY 722 CAGTTCTTACTCTAAAAACTTTGGCGCTGTACAAAGAGCGTGTGGCGCTTGTACTCTGG 781
DB 806 CCAGCTCTTATTCAAAAAACTTCGGCATGTACAAAGAGCGGTGTCGGTGCATCTTCACTTTGG 865

```



```

; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; APPLICANT: WEINER, DAVID P.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; TITLE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0390
; CURRENT APPLICATION NUMBER: US/10/979,821
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 31
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-979-821-31

Query March          21.4%; Score 255.4; DB 8; Length 1194;
Best Local Similarity 51.5%; Pred. No. 9.2e-73;
Matches 614; Conservative 0; Mismatches 576; Indels 3; Gaps 1

Qy      2   TGTTTTGAGAACATTACCGCCGCTCTGCGGACC CGCATTC TGCGGCCTGGCCGATCTGTTC 61
Db              |||||
Qy      2   TGTTTC A A A A G T T G A C G C C T A C G T G G G A C C C G N T T C T A C G C T T A T G N G C G T T T A 61
Db              |||||

Qy     62   GTGCCGATGAACGTCCCGGCAAAATAA C CT CGGGA TTGGTGTCTTA AAGANTGAGACGG 121
Db              |||||

Qy     62   AAGAAGACCCCTCGCAGCACAAAGTGAATTTAAGTATCGGTCTGTACTACAACGAAGACG 121
Db              |||||

Qy    122   GC AAA ACCCGGTACTGACGAGCTGAAAAGGCTGAACAGTATCTGCTCGNAAA--TG 178
Db              |||||

Qy    122   GAATTAATTCACAACACTCAAGCCGTCGCGAGGGGGAAGCGCGCTGAATGCGACGCTC 181
Db              |||||

Qy    179   AAACCCACAAAAATTTACCTCGGCATTGACGGCATCCCTGAA TTTGGTCGCTGCACCTCAGG 238
Db              |||||

Qy    182   ATGGCGCTTCGCTTTATTTACCGATGGAAGGCTTA ACTGCTATGCCATGCCATGCGC 241
Db              |||||

Qy    239   AAC T G C T G T T T G T A A A G G T A G C G C C T G A T C A A T G A C A A A C G T C T C G A C G G C A C A G A 298
Db              |||||

Qy    242   CGCTGCTGTTTTGGTGGCGACCATCCGGTACTGMAACAACAGCGCGTAGCAACCAATTCAA 301
Db              |||||

Qy    299   CTCGGGGGGCACTGGCGCAGCTACGCGTGGCTGCGCGATTTCTGTGCAAAAATAACACAGG 358
Db              |||||

Qy    302   CCCTTGGCGGCTCCGGGCGATTGAAGAGTGGCGCGGATTTCTGTGAACAGCTACTTCCC GG 361
Db              |||||

Qy    359   TTAAGCGTGTGGGTGAGCAACCAACGCTGSCGCAACCATGAAGAGCGTCTTTAACTCTG 418
Db              |||||

Qy    362   AATCAGGCGTCTGGGTGAGCGATCTCACTGGAAAAACCGGTAGCAATATTTCGCGGGG 421
Db              |||||

Qy    419   CAGGTC T G G A A G T T C G T G A A T A C G C T T A T T A T G A T G C G G A A A A T C A C A C T C T T G A C T T C G 478
Db              |||||

Qy    422   CTGAATCGAAGTGNATTA T T A C C C T G G T A T A C A G A G C A C T A A C G C G T G C G C T T A 481
Db              |||||

Qy    479   ATGCACTGAATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGTA GTGCTGTTC CATGGCT 538
Db              |||||

Qy    482   ATGACCTGTGGCGGACGCTGA AAAACATTA C T A C T G C C C G C A G T A T T G T G T G C T G C A T C C A T 541
Db              |||||

Qy    539   GCTGCCATAACCCACCGGTATCGACCCCTACGCTGGAACAATGGCAAACTGSCACACAC 598
Db              |||||

Qy    542   GTTGCCACAACCCACCGGTGCCGATCTCACTAATGATCAGTGGGATGCGGTGATTGAAA 601
Db              |||||

Qy    599   TCTCGGTTGAGAAAGGCTGGTTACCGCTGTTTGACTTCGCTTACACAGGGTGTTCGCCGTG 658
Db              |||||

Qy    602   TTCTAAAGCCCGGAGCTTA T T C A T T C C T C G A T A T T G C C T A T C A G G A T T G T G C G C 661
Db              |||||

Qy    659   GTCTGGAAGAGATGCTGAAGGACGTGCGCGCTTTTCGCGCTATGCAATAAAGAGCTGATTG 718
Db              |||||

Qy    662   GTATGGAAGAGATGSCCTACGCTA T T C G C G C A T T G C A G C G C T G G A T T A C C G C T C T G G 721
Db              |||||

```

		Publication No. US20050244937A1	
		GENERAL INFORMATION:	
		APPLICANT: ABRAHAM, TIMOTHY W.	
		APPLICANT: CAMERON, DOUGLAS C.	
		APPLICANT: HICKS, PAULA M.	
		APPLICANT: MCFARLAN, SARA C.	
		APPLICANT: MILLIS, JIM	
		APPLICANT: ROSAZZA, JACK	
		APPLICANT: ZHAO, LISHAN	
		APPLICANT: WEINER, DAVID P.	
		TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE	
		PRODUCTION OF MONATIN AND ITS PRECURSORS	
		FILE REFERENCE: 023829-0390	
		CURRENT APPLICATION NUMBER: US/10/979,821	
		CURRENT FILING DATE: 2004-11-03	
		PRIOR APPLICATION NUMBER: 10/422,366	
		PRIOR FILING DATE: 2003-04-23	
		PRIOR APPLICATION NUMBER: 60/374,831	
		PRIOR FILING DATE: 2002-04-23	
		NUMBER OF SEQ ID NOS: 77	
		SOFTWARE: PatentIn Ver. 3.3	
		SEQ ID NO 5	
		LENGTH: 1260	
		TYPE: DNA	
		ORGANISM: Rhodobacter sphaeroides	
		US-10-979-821-5	
		Query Match 20.3%; Score 241.6; DB 8; Length 1260;	
		Best Local Similarity 51.8%; Pred. No. 3.1e-68;	
		Matches 603; Conservative: 0; Mismatches 549; Indels 12; Gaps 2;	
Qy	25	CTGTCCGACCCGATTTCTGGGCTTGGCGGATCTGTTTCTGTCGCATGAAAGCTCCCGCAAA	84
Db	100	CCGCGGACAAAGATCTCGCACTGATCCAGATGTTTCCGCGAGGATGCGCGCGGACAA	159
Qy	85	ATTAACCTCGGATTTGGTGTCTATAAAGATGAGCGGGCAAAACCCCGTACTGACCAGC	144
Db	160	ATCGATCTGGGCTGGGCGTCTACAAGGACCCGACCGGGCTCACCCCGGTATGCGGGC	219
Qy	145	GTGAAAAGGCTGAACAGATATCTGCTCGAAAAATGAACACCAACCAAAATTTACCTCGGCA	204
Db	220	GTGAAGCGCCGAGAAAGCGGCTCTGGGAGGTCGAGACACCAAGACCTTACACCGGCTT	279
Qy	205	GACGGCATCCTGAATTTGGTGTGCTGCACTCAGGAACTGCTGTTTGTGTAAGGTAGCGC	264
Db	280	GCGGCGAGCCGCTTACAAATGCGCGATGCGGCAAGCTGATCTCG-----CAAGC	330
Qy	265	CTGATCAATGACAAACGTCGTCGACGCGACAGACTCGGGGGGCACTGGCGCACTACGC	324
Db	331	CGGTCGCGCGACCGGGTGGCTCTGGTCGCCA CCCC GCGCGGACGCGGCGCGGTGCGT	390
Qy	325	GTGGCTGCCGATTTCTTGGCAAAAAATACACGCGTTAAGCGTGTGTGGGTGAGCAACCA	384
Db	391	CAGCGCTCGAGCTGATCGCATGGCTCGCCCGAGGCGACCTGTCTGGATCTCGAAACCG	450
Qy	385	AGCTGGCCGAAACCAATGAGCGCTCTTAACTCTGACGCTCGGAAGTTGCTGAAATACGCT	444
Db	451	ACCTGGCCGAAACCATCTGTCGATCTGAAATATCTCGGCATCCCGATCGGGAATACGCG	510
Qy	445	TATTATGATCGGAAAAATCACACTCTTGACTTTGATGCACTGATTAACAGCTGAATGAA	504
Db	511	TATTTGACCGCGAGACCGGCGCGTCGATGCGGAGGCTTGATGGAGGATCTGCGCCAG	570
Qy	505	GCTCAGGCTGGCGAGTGTGTGTTCCATGCTGTGCGCAATAACCCCAACCGGTATCGAC	564
Db	571	GTGAAGCGGGGACGTTGGTGTGCTGTCACGCGTGTCTGCCAACACCCGACCGCGCCAAC	630
Qy	565	CCTACGCTGGAACAAATGGCAACA CTGGCACAACTCTCGTTGGAAGGCTGTTTACCG	624
Db	631	CCGAACCCGCTGACGTGGCTGGCGCTGTGCGAGAGCTGGCCCGGACAGGCGGTGCGC	690
Qy	625	CTGTTTGACTTCGCTTACCAAGGTTTGGCCGTGTCTGGAAGAAGATGCTGAAGGACTG	684

ESULT 6
S-10-979-821-5
Sequence 5, Application US/10979821

Db 691 CTGATCGACCTCGCCTATCAGGGCTTCGGCGACGGGCTCGAGATGGATCGGGCGGACG 750
 Qy 685 CGCGCTTCCGGCTATGCAATAAGAGCTGATTTGTTGCCAGTTCTCTACTCTAAAACTTTT 744
 Db 751 CGGCTTCGGCCACACAGACTCGCCGAGGTCTGATCGCGGCTCTGCTCGAAGAACTTC 810
 Qy 745 GGCTGTACAAACAGCGGTGTGGGGCTTGACTCTGTTGCTCGACAGTGAACCGTT 804
 Db 811 GGCATCTACCGCGAGCGAACCGGCATCTCTGA---TCGCCATCGCGGAGCGCGGGCCGG 867
 Qy 805 GATCGGCAFTTACGCCAATAAGAGCGGATTCGGGCTAACTACTCTAACCCACACAGCA 864
 Db 868 GGCACGGTGCAGGCCAATCTCACTTCTTGAACCGGCGAGAACTACTCTTCCCGCGGAC 927
 Qy 865 CACGGCGCTTCTGTTGTGCCACATCTCTGAGCAACAGATGCGTTACGTGCGATTTGGAA 924
 Db 928 CATGGCGCGGCTCGTGACCATATCTCTCGAGGACGAGCGCTGAGCGCGGACTGGAAG 987
 Qy 925 CAAGAGTGAATGATCGCCAGCGTATTCAGGCTATTCAGGCTATGCGTCAATGCTGCAATAG 984
 Db 988 GCGGAATCTCGAGAGGTGCGGCTCAACATGCTGACGCTGCGCGCCAGCTTGCAGTGG 1047
 Qy 985 CTCAGGAAAGAGCGCAACCGGCTTTCAGCTTATCATCAACAGAACGCGCATGTTTC 1044
 Db 1048 CTCAGGCCAGACCGGCTCGAACCGCTTCGGCTTCGTGGCCGAGCATCGCGGCTATGTT 1107
 Qy 1045 TCCTTCAGTGGCTGACAAAGAACCAAGTCTGCTGCGCAAGAGTTTGGCGTATAT 1104
 Db 1108 TCGCGCTCGGATCACGCCCGCGAGGTGGAGCGGCTGCGGACGAGCATCGCGGCTCTAC 1167
 Qy 1105 GCGGTTGCTTCTGTCGCTAAATGTGGCCGGGATGACACAGATGCGTCCGCTG 1164
 Db 1168 ATGGTGGCGATTCGCGGCTGAAACATCGCGGGCTGAAACCGGACGACCGCTGCGGTG 1227
 Qy 1165 TGGCAAGCGATTGTGGCAGTGTG 1188
 Db 1228 GCGCGCGGTGGCAAGGTGCTG 1251

RESULT 7
 US-11-114-922-5
 ; Sequence 5, Application US/11114922
 ; Publication No. US20050282260A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HICKS, PAULA M.
 ; APPLICANT: MCFARLAN, SARA C.
 ; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
 ; TITLE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS
 ; FILE REFERENCE: 023829-0396
 ; CURRENT APPLICATION NUMBER: US/11/114,922
 ; PRIOR FILING DATE: 2005-04-26
 ; PRIOR FILING DATE: 10/422,366
 ; PRIOR FILING DATE: 2003-04-23
 ; PRIOR FILING DATE: 60/374,831
 ; NUMBER OF SEQ ID NOS: 91
 ; SOFTWARE: PatentIn Ver. 3.3
 ; SEQ ID NO 5
 ; LENGTH: 1260
 ; TYPE: DNA
 ; ORGANISM: Rhodobacter sphaeroides
 ; US-11-114-922-5

Query Match 20.3%; Score 241.6; DB 12; Length 1260;
 Best Local Similarity 51.8%; Pred. No. 3.1e-68;
 Matches 603; Conservative 0; Mismatches 549; Indels 12; Gaps 2;
 Qy 25 CTTGCCACCCGATTTCTGGGCTGGCGGATCTGTTCTGTCGCCGATGAACGTCCGGGAAA 84
 Db 100 CCGCGGACAAAGATCTCTGCAACTGATCCAGATGTTCCGCGAGGATGCGCGGCGGACAAG 159
 Qy 85 ATTAACCTCGGATTTGTTGTTCTAATAGATGAGCGGCAAAACCCCGGTACTGACGAC 144

Db 160 ATCGATCTGGCGTGGCGCTTACAAGGACCCGACCGGGCTCACCCCGGCTCATGCGGGCC 219
 Qy 145 GTGAAAAGGCTGAACAGTATCTGCTCGAAAATGAACACCAAAAAATTTACCTCGGCATT 204
 Db 220 GTGAAGGCGCGGAGAGCGGCTCTGGAGGTTCGAGACCAACCAAGACCTTACACCGGCTT 279
 Qy 205 GACGGCATCCCTGAATTTGGTTCGCTGCACTCAGGAACTGCTGTTTGGTAAAGGTAGCGCC 264
 Db 280 GCGCGGAGCGCCCTTCAATGCCGGATGGCGAAGCTGATCCTCG-----CAGGC 330
 Qy 265 CTGATCAATGACAAAAGTCTCGACGCGACAGACTCCGGGGGGGCACTCGGCGCACTACGC 324
 Db 331 GCGGTCGCGCGGACCGGGTGGCTCGGTGCGCACCCCGCGGCGCACGGCGCGGTGCGT 390
 Qy 325 GTGGTTCGCGATTTCTTGGCAAAAATACAGGGTTAAGCGTGTGTTGGGTGAGCAACCCA 384
 Db 391 CAGGGCTCTGAGCTGATCCGATGCGCTCGCCCGAGGCCACTGTCTGGAATCTCGAACCCG 450
 Qy 385 AGCTGSCCGAACAATAAGAGCGTCTTTAACTCTGCAAGTCTTGGAACTTCTGTAATACGCT 444
 Db 451 ACTGGCCCGAACAATCTGTGATTCGTGAATATCTCGGATCCCGATGCGGGAATACCGC 510
 Qy 445 TATTATGATCGGAAAATCACACTCTTGACTTCGATGCACTGATTAACAGCTGAAATGAA 504
 Db 511 TATTTCGACGCGGAGACCGGCGCGCTCGATGCGAGGGCTTGATGAGGATCTGGCCAG 570
 Qy 505 GCTCAGGCTGGGACGTAAGTGTTCATGCGTGTGTCATTAACCAACCAACCGGTATGAC 564
 Db 571 GTGAAGCGGGGACGCTGCTGCTGCAACCGGCTGCTGCAACCAACCGGCGGCGCAAC 630
 Qy 565 CTTACGCTGGAACAATGGCAAAACACTGGCACAACTCTCGTTTGAGAAAGGCTGTTTACCG 624
 Db 631 CGGAACCGGTCAGTGGCTGGCGCTCTCGAGAGCTGGCCCGGACAGGCGGCGGACG 690
 Qy 625 CTGTTTGACTTCGCTTACAGGGTTTTGCCCGTGTCTGGAAGAAGATGCTGAAGACTG 684
 Db 691 CTGATCGACTCGCTATCAGGGCTTCGCGCGAGGCTCGAGATGATGCGGCGGCGACG 750
 Qy 685 CGCGCTTCGCGGCTATGATTAAGAGCTGATGTTGCGCAGTTCCTACTCTAAAACTTT 744
 Db 751 CGGCTTCGCGCCACACAGACTGCGCGAGGTGCTGATCGCGGCTCTCTGCTCGAAGAACTTC 810
 Qy 745 GCGCTGTACAAGAGCGGTGTGCGGCTTGTACTCTGTTGCTCGCGACAGTGAACCGTT 804
 Db 811 GGCATCTACCGGAGCGAAGCGGCTATCTGA---TCGCCATCGCGGAGGCGGGCGCG 867
 Qy 805 GATCGGCAATTCAGCCAAATGAAGCGGATTCGGGCTAACTACTCTAACCCACAGCA 864
 Db 868 GGCACGGTGCAGGCCAATCTCACTTCTTGAACCGGCGAGAACTACTCTTCCCGCGGAC 927
 Qy 865 CACGGCGCTTCTGTTGTTGCCACCATCTCTGAGCAACAGATGCGTTAGTGGGATTTGGGAA 924
 Db 928 CATGGCGCGGCTCGTGACCATGATCTCTCGAGGACGAGACGCTGAGCGCGGCTGGAAG 987
 Qy 925 CAAGAGTGAATGATGCGCGAGCGTATTCAGGCTATGCGTCAATGCTGTTGCTCAATAG 984
 Db 988 GCGGAACTCAGGAGGTGGGCTCAACATGCTGACGCTGGCGCGGCGGCTGCGGATGCG 1047
 Qy 985 CTGCGAGAAAAGGCGCAAAACCGGCTTTCAGCTTTATCATCAAAACAGAACGCGATGTTTC 1044
 Db 1048 CTGCGGCGGAGACCGGCTCGAACCGCTTCGGCTTCGTGGCCGAGCATCGCGGCTATGTT 1107
 Qy 1045 TCCTTCAGTGGCTGACAAAAGAACAAAGTGTGCGTCTGGCGGAGAGTTTGGGCTATAT 1104
 Db 1108 TCGCGCTCGGATCACGCCCGCGAGGTGGAGCGGCTGCGGACGAGCATCGCGGCTCTAC 1167
 Qy 1105 GCGGTTGCTTCTGCTGCGTAAATGTGGCGGATGACACAGATGCGTCCGCTG 1164
 Db 1168 ATGGTGGCGATTCGCGGCTGAAACATCGCGGGCTGAAACCGGACGACCGCTGCGGTG 1227
 Qy 1165 TGGCAAGCGATTGTGGCAGTGTG 1188
 Db 1228 GCGCGCGGTGGCAAGGTGCTG 1251

RESULT 8

US-10-979-821-3
; Sequence 3, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.
; APPLICANT: HICKS, PAULA M. C.
; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; APPLICANT: WEINER, DAVID P.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0390
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 3
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Rhodobacter sphaeroides
US-10-979-821-3

Query Match 20.2%; Score 240; DB 8; Length 1260;
Best Local Similarity 51.7%; Pred. No. 1e-67;
Matches 602; Conservative 0; Mismatches 550; Indels 12; Gaps 2;

QY	25	CCTGCCACCGGATCTTGGCGCTGGCGATCTGTTTCGCGCATGAACGTCGCGGCAAA	84
DB	100	CCGCGGACAAATCTTGCACACTGATCCAGATGTTCCGCGAGGATGGCGCGGACAG	159
QY	85	ATTAACCTCGGGATTTGTTGTTTAAAGATGAGACGGGCAAAACCCCGGTACTGACGAC	144
DB	160	ATCGATCTGGCGTGGCGCTTACAAGGACCGGACCGGCTCACCCCGGCTATGCGGGCC	219
QY	145	GTGAAAAGCTGAACAGTATCTCTCGAAATGAACACCAACCAATTAACCTCGGCATT	204
DB	220	GTGAAGCGCGGAGAAACGGCTCTGGGAGGTTCGAGACCAACCAAGACTTACACCGGCTT	279
QY	205	GACGGCATCTCGAATTTGGTGGTCACTCAGGAAGTCTGTTTGGTAAAGGTAGCGCC	264
DB	280	GCGGACGCGCGCTCAATGCGCGATGCGGATGGGAGTGTCTCTCG-----CGGGC	330
QY	265	CTGATCAATGACAAACGTGTGCGACGGGACAGACTCCGGGGGCGACTGGCGCACTACGC	324
DB	331	CGGTCCTCCGCGACCGGGTGGCTCGGTGCGCACCCCGCGCGACCGGCGGTGCGT	390
QY	325	GTGGTCCCATTTCTGGCAAAATATACAGCGTTAAGCGTGTGGGTGAGCAACCCA	384
DB	391	CAGGCGCTCAGCTGATCCGCGATGGCTCGCCCGAGGCCACCGCTCTGGATCTCGAACCCG	450
QY	385	AGCTGGCGCAACATAGAGCGTCTTTAACTTCGAGGTCTGGAAGTTGCTGAATAGCT	444
DB	451	ACTGGCGCAACCATCTGTCTGATCGTGAATATCTCGGCATCCGATCGGGATACCGC	510
QY	445	TATTATGATCGGAAATACACTCTTTGACTTCGATGCACTGATTAACAGCCCTGAATGAA	504
DB	511	TATTTGACCCGAGACCGCGCGTGCATGCGGAGGCGATGATGGAGGATCTGGCCAG	570
QY	505	GCTCAGGCTGGGACAGTAGTGTGTTCCATGGCTGTGCTGCAATACCAACCGGTATGAC	564
DB	571	GTGAAGCGCGGACGCTGTGCTGTGCAAGGCTGTGCTGCAACACCGGCGGCGCAAC	630
QY	565	CCTACGCTGGAAACAAATGGCAACACTTGGCACAACTCTCCGTTGAGAAAGCTGTTACCG	624

DB	631	CCGAAACCGGTGCAGTGGCTGGCCATCTGCGAGAGCCTGCGCCCGGACAGCGCGGTGCGC	690
QY	625	CTGTTTGACTTCGCTTACAGAGGTTTGGCCGTTGCTGGAAGAGAGTCTGAAGGACTG	684
DB	691	CTGATCGACCTCGCTATCAGGGCTTCGGCGACGCGGCTCGAGATGAGATGCGGCGCGACG	750
QY	685	CGGCTTTTCGCGCTATGATGATAAGAGCTGATTGTTGCCAGTTTCTTACTCTTAAAACTTT	744
DB	751	CGGCTTCGCGCACAGACTGCCCGAGGTGCTGATCGGGCTCTCTGTCGAAGACTTC	810
QY	745	GGCTGTACAAAGAGGTTTGGCGTTTACTCTCTGTTGCTGCGGACAGTGAACCGTT	804
DB	811	GGCATCTACCGGAGCGACGGCATCTCTGA---TCGCCATCGCGAGGCGGCGCGCG	867
QY	805	GATCGCGCATTCAGCCAAATGAAAGCGGCGATTCGCGCTAACTACTCTAACCCACCGACA	864
DB	868	GGCACGGTGCAGGCCAACCTCAACTTCTGAACCGGCGAGAACTACTCTTCCCGCGGAC	927
QY	865	CACGGCGCTTCTGTTGTGACCATCTGAGCAACGATGCGTTTACGTGCGATTGGGAA	924
DB	928	CATGGCGCGGCTCGTGACCATGATCTCGAGGACGAGCGCTGACGCGCGACTGGAAG	987
QY	925	CAAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGTATGCGTCAAGTTGTTGTTCAATACG	984
DB	988	GCGGAACTCGAGGAGGTGGGCTCAACATGCTGACACTGCGCGCGGAGCTTGCCTGATGCG	1047
QY	985	CTCAGGAAAAGCGGCAACCGGACTTCAGCTTTATCATCAACAGAGACGGCATGTTT	1044
DB	1048	CTCAGGCGGAGACCGGCTCGAAACCGCTTCGCGCTTCGTGGCCGAGCATCGCGGCATGTT	1107
QY	1045	TCCTTCAGTGGCTGACAAAAGAAACAGTCTGCTGTCGCGAAGAGTTTGGCGTATAT	1104
DB	1108	TCGCGCTCGGATCACCGCGCGGAGTGGAGGGCTGCGGACCGGACGCGGCTCTAC	1167
QY	1105	CGGTTGCTTCTCGTCCGTAATGTGGCGGATGACACAGATAACATGGCTCCCGCTG	1164
DB	1168	ATGTTGGCGATTCGCGGCTGAACATCGCGGGCTGAACCGGACCGCTGCGCGTCTG	1227
QY	1165	TGCGAAGCATTTGGCGAGTGTG 1188	
DB	1228	GCGCGCGGTGGCCAAAGTGTG 1251	

RESULT 9

US-11-114-922-3
; Sequence 3, Application US/11114922
; Publication No. US20050282260A1
; GENERAL INFORMATION:
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; CURRENT FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 3
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Rhodobacter sphaeroides
US-11-114-922-3

Query Match 20.2%; Score 240; DB 12; Length 1260;
Best Local Similarity 51.7%; Pred. No. 1e-67;
Matches 602; Conservative 0; Mismatches 550; Indels 12; Gaps 2;

QY	25	CCTGCCACCGGATCTTGGCGCTGGCGATCTGTTTGGTGAAGAGCTGTTACCG	84
----	----	------------------------------------------------------	----

Db 100 CCGCGGCAAGATCTCGCACTGATCCAGATGTTCCGCGAGATGCGCGCGGCAAG 159
 Qy 85 ATTAACCTCGGGATGTGTCTATAAGATGAGACGGGCAAAACCCCGGTACTGACCAAG 144
 Db 160 ATCGATCTGGCGGTGGCGGTCTACAAGACCCGACCGGGTCAACCCGGTCTATGGCGGC 219
 Qy 145 GTGAAAGGCTGACAGTATCTGCTGAAATGAAACCAACCAAAATTAACCTCGGCAAT 204
 Db 220 GTGAAGCGCGGAGAGCGGCTCTGGAGGTTCAGACCAACCAAGACTACACCGGCTT 279
 Qy 205 GACGGCATCTGAAATTTGGTGTGCTGCACTCAGGAAGTGTGTTGTTGTAAGGTAGGCGC 264
 Db 280 GCGACAGCGGCTACATGCGCGATGGGAGCTGATCTCG-----CGGC 330
 Qy 265 CTGATCAATGACAAACGTGTGCGACCGGCAAGACTCGGGGGGCACTGCGCACTACGC 324
 Db 331 CGGCTCCGCGCGACCGGTGGCTCGGTGCGCACCCCGCGGCGGCGGCGGTGCGT 390
 Qy 325 GTGCTCCGATTTCTCGGCAAAATAACAGCGTTAAGCGTGTGTGGTGAAGCAACCA 384
 Db 391 CAGCGCTCGAGCTGATCCGATGGCTCGCCCGAGGCCACCGCTCGGATCTCGAACCG 450
 Qy 385 AGCTGGCGGAACCAATAGAGCGTCTTTAACTCTGAGGTCTGGAAGTTCGTGAATAGCT 444
 Db 451 ACCTGGCGGAACCAATCTGCTGATCTGATATCTCGGATCCGATGCGGGAATACCG 510
 Qy 445 TATTATGATCGGAAATACACACTCTTTGACTTCGATGCACTGATTAACACGCTGAATGA 504
 Db 511 TATTTCGACGCGAGACCGCGCGTCTGATGCGGAGGCAATGATGAGGATCTGGCCAG 570
 Qy 505 GCTCAGGCTGGCGAGTGTGCTGTTCCATGCTGCTGCCATACCCAAACCGGTATCGAC 564
 Db 571 GTGAAGCGCGCGAGCGTGTGCTGCTGACGCGTGTGCTGCAACACCGGCGCGCAAC 630
 Qy 565 CCTACGCTGGAACCAATGGCAACACTGGCAACACTCTCCGTGGAAGGCTGCTTACCG 624
 Db 631 CCGAACCGGTGAGTGGCTGGCCATCTGCGAGAGCCTGCGCCGAGCGGCGGTGCGG 690
 Qy 625 CTGTTGACTTCGTTTACAGGCTTTTGGCGGTTTGGCGGTGCTGGAAGAGATGCTGAAGACTG 684
 Db 691 CTGATGACCTCGCTTACAGGCTTCGCGACGCGGTTCGAGATGATGATGCGGCGGACG 750
 Qy 685 CGGCTTTGGGCTATGCAATAAGAGTGTGTTGCCAGTTCCTACTCTAAAGCTTT 744
 Db 751 CGGCTTCTGGCCACCACTGCGCGAGGTGCTGATCGCGCTCTCTCTGAAAGACTTC 810
 Qy 745 GGCCTGTACACAGAGCGTGTGGCGCTTGACTCTGCTGCTGCGGACAGTGAAACCGTT 804
 Db 811 GGCATCTACCGGAGCGCACGGGATCTCTGA--TCGCCATCGCGAGGCGGCGGCGG 867
 Qy 805 GATCGCGCATTCAGCCAAATGAAGCGCGGATTCGCGCTAACTACTTAACCCACAGCA 864
 Db 868 GGCAGGTGCGAGCGCAACCTCACTCTGAAACCGGAGAACTACTCTTCCCGCGGAC 927
 Qy 865 CAGGCGGCTTCTGTTGTGCAACATCTGAGCAACGATGGTTACGTGCGATTTGGAA 924
 Db 928 CATGCGCGGCGCTGCTGACCATGATCTCGAGGACGAGACGCTGAGCGCGGCTGGAAG 987
 Qy 925 CAAGAGCTGATGATATGCGCCACGCTATTACAGGTATGCTGATGTTGCTCAATACG 984
 Db 988 CGGAACTCGAGGAGTGGGCTCAACATGCTGACATGCGCCCGGAGCTTGGCGATGCG 1047
 Qy 985 CTGAGGAAAGCGCAAAACCGGACTTCAGCTTTTATCATCAACAGAAACGGGATGTC 1044
 Db 1048 CTGAGGCGGAGACCGGCTCGAAACCGCTTCGGCTTCGTGGCGGAGCATCGCGCATGTC 1107
 Qy 1045 TCCTTCAGTGGCTGACAAAGAAACAAGTGTGCTGCTGCGCGAAGAGTTTGGGCTATAT 1104
 Db 1108 TCGCGCTTCGGGATCACGCCCGCGAGGTGAGCGGCTGCGGACCGGAGCGGGGTCTAC 1167
 Qy 1105 CGGCTTCTGCTGCGTAAATGTGCGCGGAGTGAACAGATACATGGCTCGCGCTG 1164

Db 1168 ATGGTGGCGGATTCGGCGCTGAAACATCGCGGGCTGAACCGGACCGGTCCGGTCTG 1227
 Qy 1165 TCGGAAGCGATTTGTGCGAGTGTG 1188
 Db 1228 GCGCGCGCGGTGGCCAAAGGTGCTG 1251
 RESULT 10
 US-11-136-527-2278
 ; Sequence 2278, Application US/11136527
 ; Publication No. US20050287570A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Mounts, William M
 ; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
 ; FILE REFERENCE: 031996-041000 (AM101086)
 ; CURRENT APPLICATION NUMBER: US/11/136,527
 ; CURRENT FILING DATE: 2005-05-25
 ; PRIOR APPLICATION NUMBER: US 60/574,294
 ; PRIOR FILING DATE: 2005-05-26
 ; NUMBER OF SEQ ID NOS: 362830
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 2278
 ; LENGTH: 2343
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; US-11-136-527-2278
 Query Match 18.6%; Score 221.6; DB 12; Length 2343;
 Best Local Similarity 50.1%; Pred. No. 1.5e-61;
 Matches 580; Conservative 9; Mismatches 563; Indels 6; Gaps 2;
 Qy 25 CTTGCCGACCGGATTCCTGGCCCTGGCGATCTGTTTCGTGCCGATGAACCTCCGGCAAA 84
 Db 207 CTTCCAGATCCCATCTCTGGAGTGACCGAWGCCCTTCAGAGAGATACCAACAGCAMGAAG 266
 Qy 85 ATTAACCTCGGATTTGGTGTCTATAAGATGACAGCGGCAAAACCCCGGTACTGACCAAGC 144
 Db 267 ATGAACCTGGAGTTGGTGTGCTTACCGGAGCATTAACGGAAGCTTACGTGCTCCCAAGT 326
 Qy 145 GTGAAAGGCTGAAACAGTATCTGCTGAAATGAAACCAACCAAAATTAACCTCGGCAAT 204
 Db 327 GTTCGGAAGGACAGCGCCAGATTTGCTGGGAAATTTGGWCAAGAAATACCTACCCATC 386
 Qy 205 GACGGCATCCCTGAATTTGGTGTGCTGCACTCAGGAAGTCTGTTTGTAAAGTAGGCGCC 264
 Db 387 GGGGGAATGGCTTTTGTAAAGGTTCTGCAAGACTGGCCCTGGGCGAGAACAGCGAA 446
 Qy 265 CTGATCAATGACAAAGCTGCTCGCAGGACAGACTCCGGGGGCACTGGCGCACTACGC 324
 Db 447 GTGTTGAAAGCGCGCGGTTTGTAACTGTGCAACCAATTTCCGGGACTGGAGCTTGAGG 506
 Qy 325 GTGGCTCGCGATTTCTTGGCAAAATAACAGCGTTAAAGCG---TGTGTGGGTGAGCAAC 381
 Db 507 GTCGGAGCCAGCTTTCTGCAAGATTTTAAAGTTCAGCCGAGATGCTTTCTTCCGCAAA 566
 Qy 382 CCAAGCTGGCGCAACCAATGAGCGCTTTTAACTCTGCAAGTCTGGAAGTTCTGTAATAC 441
 Db 567 CCATCTCGGGAAACCAACAGCCCATCTTCAGGGATGCGCGCATGCAAGGTTAT 626
 Qy 442 GCTTATTATGATCGGAAATACACTCTTTGACTTCGATGCACTGATTAAACAGCTCAAT 501
 Db 627 CGATCTATGACCCAGACTTCGGGCTTGTGCTTCTGTGAGCTTAGAAGACATATCA 686
 Qy 502 GAAGCTCAGGCTGGGACGTTGCTGTTCCATGGTGTGCTGCAATACCAACCGGTATC 561
 Db 687 AAATCCAGAGCAGAGTGTCTCTCTTGTGCAAGCTGCGCTCACACCCACCGCGGTG 746
 Qy 562 GACCTTACGCTGAAACAAATGGGAAACACTGGGCAACTCTCCGTTGAGAAAGGCTGTTA 621
 Db 747 GACCCGCTCCAGAGCAGTGAAGAAATGGCGCGGTGTTGTAAGAAAAAGAAATCTCTTC 806
 Qy 622 CGGCTGTTGACTTCGCTTACCAGGGTTTTCGCCCGTGG---TCTGGAAAGAGATGCTGAA 678

Db 807 GCATTTCTTTGACATGGCCCTACCAAGGCTTTGCCAGCGCGATGGTGATGAAGACGCCCTGG 866
Qy |||||
Db 679 GGACTGCGCGCTTTCCGGCTATGCTAAGAGCTGATTTGGTCCAGTTCCTTACTCTAA 738
Qy |||||
Db 867 GCCGTGCGGCACTTCATGAGCAGGATCAATGCTGCTCTGCAATCTTATGCCAAG 926
Qy |||||
Db 739 AACTTTGGCCTGTACACGAGCGTGTGGCGCTTGTACTCTGTTGCTGCGCAGAGTGA 798
Qy |||||
Db 927 AACATGGCGCTGTACGGTGAGCGTGTGGAGCCTTCACTGTGTSTGCAAGAGTGCAG 986
Qy |||||
Db 799 ACCGTGTGTCGGCATTCAGCCAAATGAAGCGGCGATTCGGGCTAACTACTCTAACCCA 858
Qy |||||
Db 987 GAAGCCAAAGGGTGGAGTCAACAGCTGAAGATCTCTGATCCGCGCTTGTATTTCCAAACCG 1046
Qy |||||
Db 859 CCAGCACACGGCGCTTCTGTTGTCGCCACCATCTCCAGCAAGAGTGCCTTACGTGCGATT 918
Qy |||||
Db 1047 CCTCTCAATGGAGCCCGGATGCCCAACCATCTGAYKTCTSSAGACKSGGAGCAA 1106
Qy |||||
Db 919 TGGGAACAAGAGCTGACTGATATGCGCCAGCGGTATTTCAGCGTATGCGTCACTGTTGTCGTC 978
Qy |||||
Db 1107 TGGTTGACAGAGGTGAAGGCTGCTGACCGCATCATCAGATGAGGACCCAGTTGTC 1166
Qy |||||
Db 979 AATAGCTGCGAGGAAAAAGGCGCAAAACCGCGACTTTCAGCTTTATCATCAAAACAGAACCGC 1038
Qy |||||
Db 1167 TCCAACTGAAGAAAGAGGCTGCTGCCCAACTGGCAGCACATCAGCGACAGATCGC 1226
Qy |||||
Db 1039 ATGTTCTCTTCAGTGGCTGACAAAGAACAGTGTGCTGCTGCGGAGAGTTGGC 1098
Qy |||||
Db 1227 ATGTTCTCTTCACCGGCTAAAGCCTGAGCAGGTGGAGCGGTGACCAAGGAGTCTTGA 1286
Qy |||||
Db 1099 GTATATGCGGTGCTTCTGCTGCGTAAATGTTGGCGGGATGACACCAAGATAAATGCT 1158
Qy |||||
Db 1287 GKCTACATGACAAAGATGCTGCTGCTCTGKGGCGGGTCACTCTGGCANATGTCGGC 1346
Qy |||||
Db 1159 CCGCTGTGCGAAGCATT 1176
Qy |||||
Db 1347 TAMCTGGCCCGCCATT 1364

RESULT 11
US-10-979-821-1
; Sequence 1, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; APPLICANT: WEINER, DAVID P.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; TITLE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0390
; CURRENT APPLICATION NUMBER: US/10/979,821
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patentin Ver. 3.3
; SEQ ID NO 1
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Sinorhizobium meliloti
US-10-979-821-1
Query Match 17.3%; Score 206.2; DB 8; Length 1170;
Best Local Similarity 50.8%; Pred. No. 1.1e-56;
Matches 580; Conservative 0; Mismatches 543; Indels 18; Gaps 3;

Qy 1 ATGTTTGAGAACATTACCGCGCTCCTGCCAGCCGATTCCTGGGCTCGCGGATCTGTTT 60
Db |||||
Qy 1 ATGTTTGAGAGCCCTCGCCCGCCAAAGCCGACCATCCCTTGTCTTTCTCTGATCGGCTGTC 60
Db |||||
Qy 61 CGTGGCGATGAAGTCCCGGCAAAATTAACCTCGGATGTTGTGCTATAAAGATGAGAGC 120
Db |||||
Qy 61 AGGAAGGATGAGCGCCCGGAAAGGTGATCTCGGCGTAGGAGTCTATCGGACGAGACC 120
Db |||||
Qy 121 GGCAAAACCCCGGTACTGACAGCGTGAAGAGCTGAACAGTATCTGCTCGAAATGAA 180
Db |||||
Qy 121 GGACGACGCCGATCTTCGGGCGGTCAAGCGCGGAAAGCGGCTTCTCGAAACACAG 180
Db |||||
Qy 181 ACCACCAAAATTAACCTCGGCAATTCAGCGCATTCCTGAAATTTGTCGTCGACTCAGGAA 240
Db |||||
Qy 181 GACAGCAAGGCTATATCGGCCCCGAGGGGACCTCGTCTTTCTCGATCGGCTCTGGGAA 240
Db |||||
Qy 241 CTGCTGTTTGGTAAAGGTAGCGCCTGATCAATGAACAAACGTGTCGACGCGACAGACT 300
Db |||||
Qy 241 CTCGT-----CGGCGCGACACGATCGAGCGGAGCCATGTTGCGGGGCTCCAGAGC 291
Db |||||
Qy 301 CCGGGGGCACTGGGCGCACTAGCGTGGCTGCGGATTTCTTGGGCAAAATACACAGCGTT 360
Db |||||
Qy 292 CCCGGCGGCTCCGGCGGCTCCGTTTGGCGCGGACCTCATCGC---CCGATGGCGGCG 348
Db |||||
Qy 361 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGCAACATAAGAGCGTCTTTAACTCTGCA 420
Db |||||
Qy 349 CGAGCATCTGCTCGGCTCGGCTGCCAGTGGCGGAAACACGCGCGGATCTTCAAGCGGCG 408
Db |||||
Qy 421 GGTCTGGAAGTTTGTGTAATACGCTTATATGATCGGAAATACACACTTTGACTTCGAT 480
Db |||||
Qy 409 GGGCTCGATATCGCACCTACGACTTCTTCGACATTCGTCGAGTCGTCATCTTCAT 468
Db |||||
Qy 481 GCATGATTAACACCTGAATGAAGCTCAGGCTGGGACGCTAGTGTCTTCATGGCTGC 540
Db |||||
Qy 469 AATCTGGTGAAGCGCTGGAAGCGCGCATCCGGCGATGCGGTGCTGCTCATGCAAGC 528
Db |||||
Qy 541 TGCCATAACCCAAACCGGTATCGACCTTAGCGTGAAACAATGGCAAAACATCGGCACAACTC 600
Db |||||
Qy 529 TGCCACAACCCGACCGCGGGCTCTGAGCGAAGCACATGGATGGAGATCGCGCGCTG 588
Db |||||
Qy 601 TCCGTTGAGAAAGGCTGTTACCGCTGTTGACTTTCGCTTACAGGGGTTTTGCCGCTG 660
Db |||||
Qy 589 GTGCGCGAGCGCGCTGCTGCGCTCGTCTGCTATCGCTATCAGGGGTTCCGGCGCGC 648
Db |||||
Qy 661 CTGGAAGAAGATGCTGAAGGACTGCGGCTTTCGCGGCTATGCAATAAGAGCTGATGTT 720
Db |||||
Qy 649 CTCGACCAAGATGTGCGGGGCTCCGGCATCTTCTCGCGGTGTCCTCCGGAAGCGCTCGTC 708
Db |||||
Qy 721 GCCAGTTTCTTCTCTAAACCTTTTGGCTGTACAAACGAGCGTGTGGGCTTGTACTCTG 780
Db |||||
Qy 709 CGGTTTCTGCTCGAAGTCTTCGGGCTTTATCGGAGCGCGCGGCGGCTATCTTCGG 768
Db |||||
Qy 781 GTTCTGCCGACAGTGAACCCGTTGATCGGCTATTCAGCCAAATGAAGCGCGGATTCGC 840
Db |||||
Qy 769 CGGACCACTCGACTGCTCGCGGACAGGGTGGCTCAAACTCGCGGGCTCGCACGC 828
Db |||||
Qy 841 GCTAACTACTTAACCCACAGCACAGGGCGCTTCTGTTGTTGGCCACCATCTGAGCAAC 900
Db |||||
Qy 829 ACCAGCTATTCTCATGCGCGCGGATCAAGCGGCGAGCGCTGTCGCGGACGATCTTTCGAC 888
Db |||||
Qy 901 GATCGGTTAGTGGGATTTGGGAACAAGAGTGAATGATGCGCCAGCGTATTCAGCGT 960
Db |||||
Qy 889 CCGGAACCTCAGGCGCGACTGGAACGAGGAGCTCGAGACGATGCGGCTCAGGATGACGGGC 948
Db |||||
Qy 961 ATGCTCAGTTGTTGCTCAATACGCTGCAAGAAAGCGCAACCCGCACTTCAGCTTT 1020
Db |||||
Qy 949 CTCGGCGGTGCTTGGCGAGGACTCCGCAACCGCTGGCAGAGC-----TCGGCGCA 1002
Db |||||
Qy 1021 ATCATCAACAGAACCGGCTGTTCTCTCAGTGGCTGACAAAGAACAAAGTGTGCGT 1080
Db |||||
Qy 1003 GTCCCGATCAGGAGGCGATGTTCTCCATGCTGCGGCTTTCCGAAGCGGAGGTTATGCGG 1062
Db |||||

Qy 1081 CTGCGCGAAGAGTTTGGGTATATGTCGGTGTCTTCTGTCGCGTAAATGTGGCGGGATG 1140
Db 1063 CTCAGGACCGAGACGGGATCTATATGCGCGCATCCGGCGCATCAACATGCGCGGGCTG 1122
Qy 1141 A 1141
Db 1123 A 1123
RESULT 12
US-11-114-922-1
; Sequence 1, Application US/11114922
; Publication No. US20050282260A1
; GENERAL INFORMATION:
; APPLICANT: HICKS, PAULA M.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; CURRENT FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 1
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Sinorhizobium meliloti
US-11-114-922-1

Query Match 17.3%; Score 206.2; DB 12; Length 1170;
Best Local Similarity 50.8%; Pred. No. 1.1e-56;
Matches 580; Conservative 0; Mismatches 543; Indels 18; Gaps 3;
Qy 1 ATGTTTGAACATACCGCGCTCTGCGACCGCATTTCTGGGCCCTGGCGGATCTGTTT 60
Db 1 ATGTTTGAACATACCGCGCTCTGCGACCGCATTTCTGGGCCCTGGCGGATCTGTTT 60
Qy 61 CGTCCGATGAACCTCCCGGCAAAATTAACCTCGGGATTGGTGTCTATAAGATGACAGC 120
Db 61 AGGAAGATGACGCCCGGCAAAAGTGTGATCTCGCGTAGAGTCTATCGGACGACCC 120
Qy 121 GCGAAAAACCCGGTACTGACCGGTGAAAAAGGCTGAACAGTATCTGCTCGAAAAATGAA 180
Db 121 GGACGACCGCGATCTTCCGGCGCTCAAGCGCGGAAAAAGCGGCTTCTCGAAAAACAG 180
Qy 181 ACCACCAAAATTACCTCGGATGACGGCATTCCTGAATTGTCGTCGTCGACTCAGAA 240
Db 181 GACAGCAAGGCTATATCGGCCCGCAAGGGGACCTCGTCTTTCTCGATCGGCTCTGGAA 240
Qy 241 CTGCTGTTGTAAAGGTAGCGCTCTGATCAATGACAAACGTCTCGCACGGACAGACT 300
Db 241 CTCGT-----CGGCGGACACAGATCGAGCGACCATGTTGGGGGTCCAGACG 291
Qy 301 CCGGGGGGCACTGGCGCACTACGGGTGCTGCGGATTTCTTGGCAAAAAATACACGGGTT 360
Db 292 CCGGGGCTCCGGCGGCTCGGTTTGGCGCGACCTCATCGC--CCGATGGCGGC 348
Qy 361 AAGCGTGTGGGTAGCAACCCAGCTGGCGCAACATAGACGCTCTTTAACTCTGCA 420
Db 349 CGAGGCATCTGGCTCGGCTCGGAGTGGCGCAACCAACCGCGCGATCTTCAAGCGGCC 408
Qy 421 GGTCTGGAAGTCTGTAATAGCTTATATGATCGGGAATACACTCTTGACTTCGAT 480
Db 409 GGGCTCGATATCGGCACCTACGACTTCTTCGACATTCGTCGACGTGGTCACTTCGAT 468
Qy 481 GCACGTGTTAAGCCTGAATGAAGAGTCAGGCTGGCGACGTAGTGTCTTCCATGGCTGC 540
Db 469 AATCTGGTGGCGGCTGGAGAGCGCCGCATCCGCGCATCGGTGCTGCTGCTGCAAGC 528

Qy 541 TGCCATAACCCAAACCGGTATCGACCCCTACGCTGAAACAAATGGCAAAACACATGGCACAACTC 600
Db 529 TGCCACAACCCGACCGCGGGCTCTGAGCGAAGCACATGGATGGAGATCGCGCGCTG 588
Qy 601 TCCGTTGAGAAAGGCTGGTTACCGCTGTTGACTTCGCTTACCCAGGTTTTGCGCGTGGT 660
Db 589 GTGGCGAGCGCGGCTGCTGCGCTCGTCACTCGCTATCAGGGGTTTCGCGCGCGC 648
Qy 661 CTGGAAGAAGATGCTGAAGGACTGCGGGCTTTTCGGCGCTATGCAATAAGAGCTGATGTT 720
Db 649 CTCGACCAAGATGTCGCGGGCTCTCGGGCATCTTCTCGGGCTGTCGCGAAGCGCTGCTC 708
Qy 721 GCCAGTTCTTACTCTAAAAAATTTGGCCTGTACAACGAGCGCTTTCGCGCTTCTACTCTG 780
Db 709 GCGGTTTCTGCTCGAAGTCTTTCGCGCTTATTCGCGAGCGCGGGCGGCGATCTTCGCG 768
Qy 781 GTTGTGCGGACAGTGAACACCGTTGATCGCGCATTCAGCCAAATGAAGGGCGGATTCGC 840
Db 769 CGGACCAAGCTCGACTGCTCGCGGACAGGGTTCGCTCAAACTCGCGGGCTTCGACGCG 828
Qy 841 GCTAACTACTCTAACCCACGACGCGGCTTCTGTTGTCGACCATCTCTGAGCAAC 900
Db 829 ACCAGCTATTCCATGCGCGGATCACGGCGAGCGCTGTCGCGACGATCTCTTGAGCAG 888
Qy 901 GATCGTTAGTTCGCTGCGATTTGGGAAACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960
Db 889 CCGAACTCAGGCGGACCTGGAGGAGGCTCGACAGTATGCGCTCAGGATGACGGGC 948
Qy 961 ATCGCTCAGTTGTTCAATACGCTGCGAGGAAAAAGGCGCAAAACCGCGACTTCAGCTTT 1020
Db 949 CTCGGCGGTGCTTTCGCGAGGACTCCGACCCGCTGGCAGAGCC-----TCGGCGCA 1002
Qy 1021 ATCATCAACAGAACGGCATGTTCTCTTCAGTGGCTGCAAAAGACAAAGTCTGCTGCT 1080
Db 1003 GTCCGCGATCAGAGGGCATGTTCTCCATGCTGCCGCTTTCCGAGCGGAGGTATCGG 1062
Qy 1081 CTCGCGAAGAGTTTCGGGTATATGCGGTGCTTCTGTCGCTAAATGTGGCGGGATG 1140
Db 1063 CTGAGACGAGCAGCGCATCTATATGCGGCGATTCGCGCGCATCAACATCGCGGGCTG 1122
Qy 1141 A 1141
Db 1123 A 1123

RESULT 13
US-10-979-821-7
; Sequence 7, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0390
; CURRENT APPLICATION NUMBER: US/10/979,821
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 7
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Leishmania major

US-10-979-821-7

```
Query Match      12.2%; Score 145.8; DB 8; Length 1239;
Best Local Similarity 48.6%; Pred. No. 6.8e-37;
Matches 492; Conservative 0; Mismatches 512; Indels 9; Gaps 3;

Qy 64 GCCGATGACGTCGCGGCAAAATTAACCTCGGGATGTTGTTCTATAAAGATGAGACGGCG 123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
94 GCCGTCGCAAGGGCCCCAAGGCCAACTCGTCTATTGTTGCTTACCGGACGAGCAGGGC 153
Qy 124 AAAACCCCGGTACTACGACGGTGAAGAGCTGAACAGTATCTGCTGAAANTGAAC 183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
154 GTCCTCTATCCGCTACGCGTTCGCGAAGCTGAGCAGCTCTCTTGGACATGAATCTC 213
Qy 184 ACCAAAATTAACCTCGGCAATGACGGCATCCCTGAATTTGGTGTGCTGCACTCAGAACTG 243
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
214 GACTACGATGATCTACCGCTACCGGCTACCGCCCTTCATCGATGAGCGGTAAAGATT 273
Qy 244 CTGTTTGGTAAAGTACGCGCTGATCAATGACAAAGCTGCTGCAACGACACTCCG 303
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
274 ATCTACGGCAA---TACCGTCGAGCTGGAGAACCTGTTGCGGTGACAGCGCTGAGCGG 330
Qy 304 GGGGCACTCGGCGCACTACGCGTGGCTGCGCATTTCTTGGCAAAAATACACAGGCTTAAG 717
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
691 AGCCTCGACACGGACGGGTATGCTGCCCGCTGTTTGGCCCGCGGCATCGAGGTACTG 750
Qy 718 GTTCCAGTTCCTACTCTAAACCTTTGGCCCTGTACACGAGCGTGTGGCGCTTGACT 777
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
751 CTGGCGCAGTGTCTCCAGAAACATGGGCTGTACAGCGAGCGTGAGGCAACGCTGTGCG 810
Qy 778 CTGGTGTGCGGACAGTGAAACCGTTGATCGCGATTCAGCCAAATGAAGACGGCGATT 837
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
811 CTGCTCTCAAGACACAGACGCGCGGATGTAAAGACGCTGATGGATTCCCTGATC 870
Qy 838 CGCGCTAACTACTTAACCCACACGACACACGCGCTTCTGTTGTGGCAACCATCCTGAGC 897
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
871 CGTGAGGAGTACAGTGGCCCCCAGCCACCGGTGCCCGCTTAGCCACCTAATCCTGAGC 930
Qy 898 AACCATGGTTACGTGGCATTTGGGACACAGCTGACTGATATGCGCCAGCGTATTTCAG 957
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
931 AACCAACGAACCTGCGAAAGGAGTGGGAGGACAGAGTATACGCCATTTGGCAGAGCGCATCCGT 990
Qy 958 CGTATCGCTCAGTTTTCGTCAATACGCTGAGGAAAGCGCAAAACCGGACTTCAGC 1017
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
991 ACGATGCGCCGCAACGGTGTACGACGAGCTGTGCGCTGACAGACGCCGGAGCTGGGAA 1050
Qy 1018 TTTATCATCAACAGAGCGGATGTTCTCTTTCAGTGGCCTTGACAAAGAAC 1070
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1051 CATGTCAATTAACAGATTGGCATGTTTTCCTTCTCGGGCTGTCAAAGGGCGCA 1103
```

RESULT 14

```
US-11-114-922-7
; Sequence 7, Application US/11114922
; Publication No. US20050282260A1
; GENERAL INFORMATION:
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; PRIOR FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 7
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Leishmania major
; US-11-114-922-7
```

```
Query Match      12.2%; Score 145.8; DB 12; Length 1239;
Best Local Similarity 48.6%; Pred. No. 6.8e-37;
Matches 492; Conservative 0; Mismatches 512; Indels 9; Gaps 3;

Qy 64 GCCGATGACGTCGCGGCAAAATTAACCTCGGGATGTTGTTCTATAAAGATGAGACGGCG 123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
94 GCCGTCGCAAGGGCCCCAAGGCCAACTCGTCTATTGTTGCTTACCGGACGAGCAGGGC 153
Qy 124 AAAACCCCGGTACTGACGCGTGAAGAGCTGAACAGTATCTGCTGCAAAAATGAAC 183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
154 GTCCTCTATCCGCTACGCGTGGTCCGCAAGCTGAGCAGCTTCTTGGACATGAATCTC 213
Qy 184 ACCAAAATTAACCTCGGCAATGACGGCATCCCTGAATTTGGTGTGCTGCACTCAGAACTG 243
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
214 GACTACGATGATCTACCGCTACCGGCTACCGCCCTTCATCGATGAGCGGTAAAGATT 273
Qy 244 CTGTTTGGTAAAGTACGCGCTGATCAATGACAAAGCTGCTGCAACGACACTCCG 303
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
274 ATCTACGGCAA---TACCGTCGAGCTGGAGAACCTGTTGCGGTGACAGCGCTGAGCGG 330
Qy 304 GGGGCACTCGGCGCACTACGCGTGGCTGCGCATTTCTTGGCAAAAATACACAGGCTTAAG 363
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
331 ACCGCTGCTGTCTCTCTCGGGGCGAAGCTGCTGACTCGCGCTTTCGACGCTGAGACGACG 390
Qy 364 CGTGTGTGGGTGAGCAACCCCAAGCTGGCGCAACCATGAAGACGCTTTAACTCTGCAG-- 421
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
391 CCCATCTPACCTTTCCGACCCCGACGTTGGCCCAACCACTACGCGCTGCGTGAAGGCTGCTGCG 450
Qy 422 -GTCGTGAAGTTCGTGAATACGCTTATTATGATGCGGAAATACACACTCTTTGACTTCGAT 480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
451 TGGNAGAACATCTGCAGTACGCTACTACGACCCCAAGCGGTACGCTGAAATTCGAG 510
Qy 481 GCACTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGCTAGTGTGTTTCATGCTGCTG 540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
511 GGCATGAAGAAAGACATTTCTGGGGGCGCGGACGCGCTCCGCTGTTTCACTTCTGACACAGTGC 570
Qy 541 TGCATTAACCCCAACGGGTATCGACCCCTACCGTGGAAACATGGGCAACACTGGCACACTC 600
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
571 GCGCACAAACCCCGCGGTGGACCGGTGCGAGGACGATGGAAACGAGATCGCTCACTG 630
Qy 601 TCCGTTGAGAAAGGTGTTTACCCTGTTTGAATTCGCTTACCAAGGGTTTTGC---CCGT 657
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
631 ATGCTGCCAAGCACCATCAGGTGTTCTTCACTCCGCTACCAAGGCTATGCGAGCGG 690
Qy 658 GGTCTGGAAGAAGATGCTGAAGGACTGCGCGCTTTGCGGCTATGCAATAAAGACTGATT 717
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
691 AGCCTCGACACGGACGGGTATGCTGCCCGCTGTTTGGCCCGCGGCATCGAGGTACTG 750
Qy 718 GTTCCAGTTCCTACTCTAAACCTTTGGCCCTGTACACGAGCGTGTGGCGCTTGACT 777
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
751 CTGGCGCAGTGTCTCCAGAAACATGGGCTGTACAGCGAGCGTGAGGCAACGCTGTGCG 810
Qy 778 CTGGTGTGCGGACAGTGAAACCGTTGATCGCGATTCAGCCAAATGAAGACGGCGATT 837
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
811 CTGCTCTCAAGACACAGACGCGCGGATGTAAAGACGCTGATGGATTCCCTGATC 870
Qy 838 CGCGCTAACTACTTAACCCACACGACACACGCGCTTCTGTTGTGGCAACCATCCTGAGC 897
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
871 CGTGAGGAGTACAGTGGCCCCCAGCCACCGGTGCCCGCTTAGCCACCTAATCCTGAGC 930
Qy 898 AACCATGGTTACGTGGCATTTGGGACACAGCTGACTGATATGCGCCAGCGTATTTCAG 957
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
931 AACCAACGAACCTGCGAAAGGAGTGGGAGGACAGAGTATACGCCATTTGGCAGAGCGCATCCGT 990
Qy 958 CGTATCGCTCAGTTTTCGTCAATACGCTGAGGAAAGCGCAAAACCGGACTTCAGC 1017
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
991 ACGATGCGCCGCAACGGTGTACGACGAGCTGTGCGCTGACAGACGCCGGAGCTGGGAA 1050
Qy 1018 TTTATCATCAACAGAGCGGATGTTCTCTTTCAGTGGCCTTGACAAAGAAC 1070
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1051 CATGTCAATTAACAGATTGGCATGTTTTCCTTCTCGGGCTGTCAAAGGGCGCA 1103
```

```
Qy 718 GTTCCAGTTCTACTTAAAACTTTGGCTGTACACGAGCGTGTGGCGTTGACT 777
Db 751 CTGGCGAGTGTGTTCTCAAGACATGAGCTTTGTACAGCGAGCGTCAGGACAGCTGTG 810
Qy 778 CTGCTTGTGTCGACAGTGAACCGTTGATCGCGCATTTACGCCAAATGAAGCGCGGATT 837
Db 811 CTGCTCTCAAGGACAAGACGCGCGGATGTAAAGAGCGTGATGCTGATC 870
Qy 838 CGCGCTAACTACTTAACCCACAGCACAGCGCGCTTCTGTTGTGGCCACCCTGAGC 897
Db 871 CGTGAGGATACAGTGTGCGCCCAAGCCAGCGGTGCGCGCTTAGCCACCTAAATCCTGAGC 930
Qy 898 AACCATGGTTACGTGGGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGPATTCAG 957
Db 931 AACAACTGCGGAAAGAGTGGGAGCGAGAGCTATCAGCCATGGCAGAGCGCATCCGT 990
Qy 958 CGTATGGTCAAGTTGTTTCGTAATACGCTGAGGAAAAAGCGCAAAACCGCGACTTCAGC 1017
Db 991 ACGATGCGCGCACCGGTGACGAGCTGCTGCGCTGCGAGCGCCGCGGAGCTGGGA 1050
Qy 1018 TTTATCATCAACAGACGCGCATGTTCTCTTTCAGTGGCTGACAAAGAAC 1070
Db 1051 CATGTCATTAACAGATTTGCGATGTTTCTTCTCTCGGGTGTCAAAGGGCCA 1103

RESULT 15
US-10-932-182A-82350
; Sequence 82350, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82350
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-82350
```

```
Query Match 12.2%; Score 145.2; DB 7; Length 1251;
Best Local Similarity 48.4%; Pred. No. 1.1e-36;
Matches 572; Conservative 0; Mismatches 583; Indels 27; Gaps 5;

Qy 2 TGTTTGAGAACATTACCGCCCTCTCTGCCGACCGGATTCTGGGCGCTGGCGGATCTGTTTC 61
Db 56 TGTTCAATAACATCGAATTGCTGCCCTCTGATGCCCTTTTGGTATTAAAGCAAGGTACG 115
Qy 62 GTGCCGATGAACGTCGCCGCAAAATTAACCTCGGGATGGTGTCTTAAGATGAGACGG 121
Db 116 GGCAAGATCAACGTGCTACCAAGGTGCACTTGGGTATCGGGGCTTACAGAGACGACAACG 175
Qy 122 GCNAACCCCGGTACTGACCGCGCTGAAAGGCTGAACAGTATCTCTCGAAATGAA 181
Db 176 GTAACCACTGGGTCTTGCAAGTGTAAAGCGCGCGAAAGCTAAITCATACACAGCT 235
Qy 182 CCAGCAAAAAT---TACCTCGGCATTGACGGCATCCCTGAAATTTGGTCGCTGCACTCAGG 238
Db 236 CTTACAACCATGAATACCTCGGTATTACCGGTCTGCCAAGTTTGACATCTAACGCCGCCA 295
Qy 239 AACTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCACGGCACA 298
Db 296 AGATCATCTTCGGTACGCAATCCGATGCTTGCAGGAAGACAGAGTAAATCTCAGTACAAT 355
Qy 299 CTCGGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTCGGCAAAAATACCAAGC 358
```

```
Db 356 CACTGTCTGGTACGGGTGCTCTTCATATATTTTCCGAAGTTTTTTTTCAAAAATCTTCCAG 415
Qy 359 TTAAGCGTGTGGGTGAGCAACCCAGCTGGCGGACCATTAAGAGCGTCTTTAACTCTG 418
Db 416 ATAACCTGGTCTATTGTTCTAAGCTTACTTTGGGCCAACACATGGGCCATTTTGGAGAAC 475
Qy 419 CAGGTCTGGAAGTTCTGTAATACGTTATATATCGGAAAAATCACACTCTTGACTTCG 478
Db 476 AAGCTTTGAJAAACGGGCACTTACCTTACTGGGCCAACGAACTAAGTCTTTGGACCTAA 535
Qy 479 ATGCACTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAAGTGTCTTCCATGGCT 538
Db 536 ACGGCTTTCTAAATGCTATTTCAAAAAGCTCCAGAGGCTCCATTTTCGTTCTGCACTCT 595
Qy 539 GCTGCCATTAACCAACGGGTATCGACCTACGCTGGAACAAATGGCAAACTGTCGACCAAC 598
Db 596 GCGCCCATTAACCAACTGGTCTGGACCTACTAGTGAACAAATGGGTTCAAAATCGTTGATG 655
Qy 599 TCTCCGTTGAGAAGCTGGTTACCGCTGTTTACCTTTCGCTTACCGGGTTTTGCGCGTG 658
Db 656 CTATCGCTCAAAAAACCATCGCTTATTTGACACCGCTACCAAGGTTTTGCCACTG 715
Qy 659 G---TCTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGCTATGCAATAAGAG--- 711
Db 716 GAGATTTGGACAAGGATGCTATGCTGCTAGTGTGGAGAAAGCTTTCGACGGTCT 775
Qy 712 -----CTGATTGTTGCCAGTTCCTACTCTTAAAACTTTGGCCTGTACACGAGCGGTG 766
Db 776 CTCGCGTCTTTGCTGCTCAGTCTCTTTGCCAAGAACCGCGTATGTACGGGTGAGCGGTAG 835
Qy 767 GCGCTTGACTCTGCTGTTGCT-----GCCACAGTGAACCGTTTGTATCGCGCATTTCA 817
Db 836 GTTGTTTCCATCTAGCACTTACAAAACAAGCTCAAACAACAACTATAAAGCCTGCTGTTA 895
Qy 818 G---CCAAATGAAGCGGCGATTCGCGCTAACTACTCTAAACCACGACACACGGGCTT 874
Db 896 CATCTCAATTTGGCCAAAATCAATTCGTAGTGAAGTGTCCAACCCACCGGCTACCGGCTA 955
Qy 875 CTGTTGTTGCCACCATCTCTGAGCAACGATGCTGAGTTCGTTGCGGATTTGGGAACAAGACTGA 934
Db 956 AGATTGTCGCTAACTGTTGGAAACGCCAGAAATTAACGGAACAGTGGCACAAGGATATGG 1015
Qy 935 CTGATATGCCACGCGTATTTCAGGTAATGCGTCAGTTGTTTCGTCATACCTGCAGGAAA 994
Db 1016 TTACCATGCTCTCCAGAAATTAAGAAATGAGGCGCGCATTTAAGAGACCATTTAGTCAAGT 1075
Qy 995 AAGCGCAAAACCCGCGACTTTCAGCTTTTATCATCAAAACAGAACCGCATGTTCTCTTCAAGT 1054
Db 1076 TGGGCACTCTGCGCACTGGGATCATATAGTAAATCAATGCGGGATGTTCTCTTTTACAG 1135
Qy 1055 GCTTGACAAAAGAACAAAGTGTGCGTCTGGCGGCAAGAGTTTGGCGTATATGCGGTTGCTT 1114
Db 1136 GATTGACTCTCTCAATGTTTAAACGACTTTGAAGAAAACCCACGCGATTTTACTTGGTTGCC 1195
Qy 1115 CTGCTCGCTTAAATGTTGGCGGATGACACCAAGTAAACATG 1156
Db 1196 CAGGTAGAGCTTCTATTGCTGGATTGAATCAAGAAACGTEG 1237
```

Search completed: March 15, 2006, 14:54:58
Job time : 1440 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 17:56:20 ; Search time 40 Seconds
(without alignments)
952.546 Million cell updates/sec

Title: US-10-673-786A-2
Perfect score: 2045
Sequence: 1 MFENITRAPADPILGLADLF.....VAGTMDNMAPLCEAIVAVL 396
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2045	100.0	396	1 XNFCDD	aspartate transami
2	2037	99.6	396	2 A85619	aspartate aminotra
3	2037	99.6	396	2 C90755	aspartate aminotra
4	1778	96.7	396	2 AD0616	aspartate aminotra
5	1775	86.8	396	2 AD0172	aspartate transami
6	1329	65.0	396	2 I64132	aspartate transami
7	1315	64.3	413	2 A82217	transaminase (EC 2
8	1311	64.1	397	2 B81915	aspartate transami
9	1310	64.1	397	2 C81188	aspartate transami
10	982	48.0	398	2 B83252	probable amino aci
11	910.5	44.5	399	2 D83535	aromatic-amino aci
12	909	44.4	400	2 C82855	aromatic-amino aci
13	909	44.4	407	2 JC5124	aspartate transami
14	908	44.4	409	2 S53303	aspartate transami
15	904	44.2	405	2 T14311	aspartate transami
16	891.5	43.6	394	2 T06197	aromatic-amino aci
17	885	43.3	456	2 T06136	aspartate transami
18	879	43.0	418	2 S46315	aspartate transami
19	874	42.7	453	2 T04646	aspartate transami
20	873.5	42.7	404	2 G82449	amino acid biosynt
21	871	42.6	449	2 T48511	aspartate transami
22	863	42.2	463	2 S33528	aspartate transami
23	861	42.1	454	1 XNVLB	aspartate transami
24	861	42.1	455	2 S46316	aspartate transami
25	861	42.1	457	2 S65675	aspartate transami
26	859.5	42.0	397	2 AD0040	aromatic-amino aci
27	843.5	41.2	397	1 XNFCY	aromatic-amino aci
28	842.5	41.2	397	2 H86098	tyrosine aminotran
29	842.5	41.2	397	2 D91258	tyrosine aminotran

30	839.5	41.1	397	2 S71928	aromatic-amino-aci
31	838.5	41.0	397	2 AG1016	aromatic-amino-aci
32	838	41.0	397	2 H81054	aromatic-amino-aci
33	832	40.7	397	2 G81821	aromatic-amino-aci
34	832	40.7	453	2 S39928	aspartate transami
35	829	40.5	453	2 S39927	aspartate transami
36	829	40.5	463	2 S39925	aspartate transami
37	827	40.4	463	2 S39926	aspartate transami
38	818	40.0	389	2 B47094	aromatic-amino-aci
39	812	39.7	430	1 XNPGDM	aspartate transami
40	809.5	39.6	449	2 T30955	probable aspartate
41	808	39.5	430	1 XNRTDM	aspartate transami
42	805	39.4	430	2 S35960	aspartate transami
43	802.5	39.2	414	2 T15494	aspartate transami
44	801	39.2	423	1 XNCHDM	aspartate transami
45	798.5	39.0	433	2 S01174	aspartate transami

ALIGNMENTS

RESULT 1
XNFCDD

aspartate transaminase (EC 2.6.1.1) aspC [validated] - Escherichia coli (strain K-12)
N:Alternate names: aspartate aminotransferase; transaminase A
C:Species: Escherichia coli
C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 09-Jul-2004
C:Accession: A00598; A38045; A29306; G64832
R:Kuramitsu, S.; Okuno, S.; Ogawa, T.; Ogawa, H.; Kagamiyama, H.
J. Biochem. 97, 1259-1262, 1985
A:Title: Aspartate aminotransferase of Escherichia coli: nucleotide sequence of the aspC
A:Reference number: A00598; MUID:85289110; PMID:3897210
A:Accession: A00598
A:Molecule type: DNA
A:Residues: 1-396 <KUR>
A:Cross-references: UNIPROT:P00509; UNIPARC:UPI000011044B; GB:X05904; NID:g41012; PIDN:AA

R:Kondo, K.; Wakabayashi, S.; Yagi, T.; Kagamiyama, H.
Biochem. Biophys. Res. Commun. 122, 62-67, 1984
A:Title: The complete amino acid sequence of aspartate aminotransferase from Escherichia
A:Reference number: A38045; MUID:84256832; PMID:6378205
A:Accession: A38045
A:Molecule type: protein
A:Residues: 1-396 <KON1>
A:Cross-references: UNIPARC:UPI000011044B
R:Kondo, K.; Wakabayashi, S.; Kagamiyama, H.
J. Biol. Chem. 262, 8648-8659, 1987
A:Title: Structural studies on aspartate aminotransferase from Escherichia coli. Covalen
A:Reference number: A29306; MUID:87250482; PMID:3298240
A:Accession: A29306
A:Molecule type: protein
A:Residues: 1-396 <KON2>
A:Cross-references: UNIPARC:UPI000011044B
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC

.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: G64832
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-396 <BLAT>
A:Cross-references: UNIPARC:UPI000011044B; GB:A5000195; GB:U00096; NID:g1787156; PIDN:AA
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: aspC
A:Map position: 21 min
C:Complex: homodimer
C:Function:
A:Description: catalyzes the reversible transfer of the amino group from L-aspartate to
A:Pathway: aspartate catabolism
A>Note: in eukaryotes, there are two isozymes: one is located in the mitochondrial matrix
C:Superfamily: aspartate aminotransferase
C:Keywords: aminotransferase; homodimer; pyridoxal phosphate

F:246/Binding site: pyridoxal phosphate (lys) (covalent) #status experimental
F:374/Binding site: substrate (Arg) #status predicted

Query Match 100.0%; Score 2045; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 9.3e-153;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MFENITAAPADPILGLADLFRADERPGKINLIGIVYKDETKTPTVLTSVKKAQYLLENE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINLIGIVYKDETKTPTVLTSVKKAQYLLENE 60

Qy 61 TTKNYLIGIDGIPFGRCTQELLFGKGSALINDKARATQTPGGTGCALRVAADFLLAKNTSV 120
Db 61 TTKNYLIGIDGIPFGRCTQELLFGKGSALINDKARATQTPGGTGCALRVAADFLLAKNTSV 120

Qy 121 KRVVSNPSPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
Db 121 KRVVSNPSPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180

Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKGLPLDFPAYQGFARGLEEDAEGLRAPAAHMKELIV 240
Db 181 CHNPTGIDPTLEQWQTLAQLSVEKGLPLDFPAYQGFARGLEEDAEGLRAPAAHMKELIV 240

Qy 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300

Qy 301 DALRAIWEQELTDMRQRIQRMQLFVNTLQEKGNRDFSFIKONGMFSFGLTKEQVLR 360
Db 301 DALRAIWEQELTDMRQRIQRMQLFVNTLQEKGNRDFSFIKONGMFSFGLTKEQVLR 360

Qy 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
```

RESULT 2

A85619
aspartate aminotransferase [imported] - Escherichia coli (strain O157:H7, substrain EDL93)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: A85619
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: A85619
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <STO>
A:Cross-references: UNIPROT:Q8XDF3; UNIPARC:UPI000000D037B; GB:AE005174; NID:g12514100; F
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: aspC
C:Superfamily: aspartate aminotransferase

Query Match 99.6%; Score 2037; DB 2; Length 396;
Best Local Similarity 99.5%; Pred. No. 4e-152;
Matches 394; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 MFENITAAPADPILGLADLFRADERPGKINLIGIVYKDETKTPTVLTSVKKAQYLLENE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINLIGIVYKDETKTPTVLTSVKKAQYLLENE 60

Qy 61 TTKNYLIGIDGIPFGRCTQELLFGKGSALINDKARATQTPGGTGCALRVAADFLLAKNTSV 120
Db 61 TTKNYLIGIDGIPFGRCTQELLFGKGSALINDKARATQTPGGTGCALRVAADFLLAKNTSV 120

Qy 121 KRVVSNPSPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
Db 121 KRVVSNPSPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
```

```
Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKGLPLDFPAYQGFARGLEEDAEGLRAPAAHMKELIV 240
Db 181 CHNPTGIDPTLEQWQTLAQLSVEKGLPLDFPAYQGFARGLEEDAEGLRAPAAHMKELIV 240

Qy 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300

Qy 301 DALRAIWEQELTDMRQRIQRMQLFVNTLQEKGNRDFSFIKONGMFSFGLTKEQVLR 360
Db 301 DALRAIWEQELTDMRQRIQRMQLFVNTLQEKGNRDFSFIKONGMFSFGLTKEQVLR 360

Qy 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
```

RESULT 3

C90755

aspartate aminotransferase [imported] - Escherichia coli (strain O157:H7, substrain RIMD
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: C90755
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Iehii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: C90755
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <HAY>
A:Cross-references: UNIPROT:Q8XDF3; UNIPARC:UPI000000D037B; GB:BA000007; PIDN:BAB34434.1;
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs1011
C:Superfamily: aspartate aminotransferase

Query Match 99.6%; Score 2037; DB 2; Length 396;
Best Local Similarity 99.5%; Pred. No. 4e-152;
Matches 394; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 MFENITAAPADPILGLADLFRADERPGKINLIGIVYKDETKTPTVLTSVKKAQYLLENE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINLIGIVYKDETKTPTVLTSVKKAQYLLENE 60

Qy 61 TTKNYLIGIDGIPFGRCTQELLFGKGSALINDKARATQTPGGTGCALRVAADFLLAKNTSV 120
Db 61 TTKNYLIGIDGIPFGRCTQELLFGKGSALINDKARATQTPGGTGCALRVAADFLLAKNTSV 120

Qy 121 KRVVSNPSPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
Db 121 KRVVSNPSPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180

Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKGLPLDFPAYQGFARGLEEDAEGLRAPAAHMKELIV 240
Db 181 CHNPTGIDPTLEQWQTLAQLSVEKGLPLDFPAYQGFARGLEEDAEGLRAPAAHMKELIV 240

Qy 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300

Qy 301 DALRAIWEQELTDMRQRIQRMQLFVNTLQEKGNRDFSFIKONGMFSFGLTKEQVLR 360
Db 301 DALRAIWEQELTDMRQRIQRMQLFVNTLQEKGNRDFSFIKONGMFSFGLTKEQVLR 360

Qy 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
```

RESULT 4

AD0616

aspartate aminotransferase [imported] - Salmonella enterica subsp. enterica serovar Typh
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A>Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AD0616
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conner, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moulie, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AD0616
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <PAR>
A:Cross-references: UNIPARC:UP1000005A061; GB:AL513382; PIDN:CAD05398.1; PID:g16502160;
C:Genetics:
A:Gene: Sry1000
C:Superfamily: aspartate aminotransferase
Query Match 96.7%; Score 1978; DB 2; Length 396;
Best Local Similarity 95.7%; Pred. No. 1.7e-147;
Matches 379; Conservative 12; Mismatches 5; Indels 0; Gaps 0;
Qy 1 MFENITAAPADPILGLADLFRADRPCKINLIGVYKDETKPTVLTTSVKAEQYLLNE 60
Db 1 MFENITAAPADPILGLADLFRADRPCKINLIGVYKDETKPTVLTTSVKAEQYLLNE 60
Qy 61 TTKNYLIGDGIPEFCRCCTOELLFGKGSALINDKARTATPGTGTALRVAADFLAKNTSV 120
Db 61 TTKNYLIGDGIPEFCRCCTOELLFGKGSALINDKARTATPGTGTALRVAADFLAKNTSV 120
Qy 121 KRNVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSINEAQAGDVVLFHGC 180
Db 121 KRNVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSINEAQAGDVVLFHGC 180
Qy 181 CHNPTGIDPTLEQWOTLAQSLVEKGMPLFDFAQGFARGLDEEGLRAFAAMHKEIV 240
Db 181 CHNPTGIDPTLEQWOTLAQSLVEKGMPLFDFAQGFARGLDEEGLRAFAAMHKEIV 240
Qy 241 ASSYKFNGLYNERVAGCTLVAADETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
Db 241 ASSYKFNGLYNERVAGCTLVAADETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
Qy 301 DALRAIWQEQLTDMQRQIRQRMQLFVNTLQKGNRDFSFIKONGMFSFGLTKQVLR 360
Db 301 DALRAIWQEQLTDMQRQIRQRMQLFVNTLQKGNRDFSFIKONGMFSFGLTKQVLR 360
Qy 361 LREEFGVYAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
Db 361 LREEFGVYAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
RESULT 5
AD0172
aspartate transaminase (EC 2.6.1.1) [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AD0172
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarrag, K.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AD0172
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <KUR>
A:Cross-references: UNIPROT:Q8ZG95; UNIPARC:UP100000CD7E2; GB:AL590842; PIDN:CAC90239.1;
C:Genetics:
A:Gene: aspc

C:Superfamily: aspartate aminotransferase
C:Keywords: aminotransferase
Query Match 86.8%; Score 1775; DB 2; Length 396;
Best Local Similarity 86.1%; Pred. No. 1.5e-131;
Matches 341; Conservative 16; Mismatches 39; Indels 0; Gaps 0;
Qy 1 MFENITAAPADPILGLADLFRADRPCKINLIGVYKDETKPTVLTTSVKAEQYLLNE 60
Db 1 MFEKITAAPADPILGLTDFRADRAHKNLIGVYKDETKPTVLTTSVKAEQYLLNE 60
Qy 61 TTKNYLIGDGIPEFCRCCTOELLFGKGSALINDKARTATPGTGTALRVAADFLAKNTSV 120
Db 61 ATKNYLIGDGLPVFASCTQELLFGANSALINDKARTATPGTGTGGLRIADFLAHTSA 120
Qy 121 KRNVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSINEAQAGDVVLFHGC 180
Db 121 KRNVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSINEAQAGDVVLFHGC 180
Qy 181 CHNPTGIDPTLEQWOTLAQSLVEKGMPLFDFAQGFARGLDEEGLRAFAAMHKEIV 240
Db 181 CHNPTGIDPTLEQWOTLAQSLVEKGMPLFDFAQGFARGLDEEGLRAFAAMHKEIV 240
Qy 241 ASSYKFNGLYNERVAGCTLVAADETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
Db 241 CSSYKFNGLYNERVAGCTLVAADETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
Qy 301 DALRAIWQEQLTDMQRQIRQRMQLFVNTLQKGNRDFSFIKONGMFSFGLTKQVLR 360
Db 301 AALRAIWQEQLTDMQRQIRQRMQLFVNTLQKGNRDFSFIKONGMFSFGLTKQVLR 360
Qy 361 LREEFGVYAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
Db 361 LREEFGVYAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
RESULT 6
IG4132
aspartate transaminase (EC 2.6.1.1) - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A>Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: IG4132
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-396 <TIGR>
A:Cross-references: UNIPROT:P44425; UNIPARC:UP100001250E7; GB:U32835; GB:L42023; NID:g15
C:Superfamily: aspartate aminotransferase
C:Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate
F:246/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted
Query Match 65.0%; Score 1329; DB 2; Length 396;
Best Local Similarity 62.6%; Pred. No. 1.5e-96;
Matches 248; Conservative 61; Mismatches 87; Indels 0; Gaps 0;
Qy 1 MFENITAAPADPILGLADLFRADRPCKINLIGVYKDETKPTVLTTSVKAEQYLLNE 60
Db 1 MFEHITAAPADPILGLGEAFKSETRENKINLIGVYKDAQGTTPIMHVAKAEKRLFDKE 60
Qy 61 TTKNYLIGDGIPEFCRCCTOELLFGKGSALINDKARTATPGTGTALRVAADFLAKNTSV 120
Db 61 KTKNYLIGDGIADYNEQTALLFGKDESVIOSNARTVQSLGGTALRIAEFIKQRTKA 120
Qy 121 KRNVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSINEAQAGDVVLFHGC 180
Db 121 QNVWISTPTWPNHNAIFNAVGMTIREYRYDAERKALDWEHLEDLSQASEGDDVLLHGC 180

[illegible]

```

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-397 <TT>
A:Cross-references: UNIPROT:Q9K0P5; UNIPARC:UPI00000C44E9; GB:AE002409; GB:AE002098; NITD
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0540
C:Superfamily: aspartate aminotransferase
C:Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate
F:247/Binding site: Pyridoxal phosphate (lys) (covalent) #status predicted

Query Match      64.1%; Score 1310; DB 2; Length 397;
Best Local Similarity 61.8%; Pred. No. 4,7e-95;
Matches 244; Conservative 64; Mismatches 87; Indels 0; Gaps 0;

Qy 2 PENTTAAPADPILGLADLFRADERPGKINIGIVYKDETGTCTPVLTSVKKAEQVLIENET 61
Db 3 FKHYEAAAPDILGLGSAFKAETPEKVNIGIVYKDSAGATPLVKAVKAEKRLLESET 62
Qy 62 TKNVYLGIDGPIPEGRCRTQELLFGKGSALINDKARTAQTPGGTGCALRVAAADFLAKNTSVK 121
Db 63 TKNVLTIDGVADYNAQTQILLFGKDHIIASRRAKTAQSLGGTGCALRIAEFAKRLQNAQ 122
Qy 122 RVVVSNSPWNHKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEPAQAGDVVLPHGCC 181
Db 123 TIWISNTPNHNHAIKAAVGIQDKPYRYDYAAKHGLDWDGMIEDLSQAQKGDIVLLHGCC 182
Qy 182 HNPYIGIDPTLEOMQTAAQLSVKQWLPFLDFAYOGFARGLEEDAEGIRAFAMHKEIIVA 241
Db 183 HNPYIGIDPTPEQWETLAKLSAEKQWLPFLDFAYOGFGNGLEEDAYGLURVFKHNTLLIA 242
Qy 242 SSKSNKFNGLYNERVGAFTLVAAADSETVDRAFSSQMKAAIRANYSNPPAHGASVVATILSND 301
Db 243 SSKSNKFGMYNERVGAFTLVAAEDDEETAARAHQSQVKTITRTLYSNPASHGANTIALVLKND 302
Qy 302 ALRAINEQEELTDARQRTQRMQLFVNTLOEKGANDRDSFLIKONGHSPFSGLTKEQVLR 361
Db 303 DLKAQWTAELDEWMGRIKAMQRKVGLLKAKGASQNFDFIIKONGMFSFSGLTPEQVDRL 362
Qy 362 REEFQGVYAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
Db 363 KRFAYVAVSGRINVAGITDNNIDYLCESIKVYI 397

```

```

RESULT 10
B83252
C:Probable amino acid aminotransferase PA3139 [imported] - Pseudomonas aeruginosa (strain
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: B83252
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B83252
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-398 <STO>
A:Cross-references: UNIPROT:P72173, UNIPARC:UPI00001250EA; GB:AE004738; GB:AE004091; NII
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3139
C:Superfamily: aspartate aminotransferase

Query Match 48.0%; Score 982; DB 2; Length 398;
Best Local Similarity 48.0%; Pred. No. 2.6e-69;
Matches 190; Conservative 70; Mismatches 136; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADRPQKINLIGVGYKDETKTPVLTSVKAEQYLENE 60
Db 3 LFGAVENAPDRPILGLNEAFNADTRPQKINLIGVGYNEEGRIPLLRVAOAAEKARIEAH 62

```

```

Qy 61 TTKNYLGIDGTFEFCRQCTQELLFGKGSALINDKRTAQTPTGGTGAURVAADFLAKNTSV 120
Db 63 APRGGLPTIEGIAAYDQGVQKLLFGNESELLAAGRVVTTQAVGGTGAUKLGADFLEKRLPD 122
Qy 121 KRVMVSNPWNHKSFVNSAGLEVREYAYDAENHTLDFDALINSLNEAOAGDVVLPHGC 180
Db 123 ATVAISDPSWENHRALFEAGPVPQNYRYDAASGNVNRAGLLEDLNALPARSIVVLHAC 182
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPLEDFAYQGFARGLEEDAEGIRAPAAHMKELIV 240
Db 183 CHNPTGVDLELDDWKQVLDLAKGHVPFLDIAYQGFNGIEEDAAAARVLFQAQSGLSFFV 242
Qy 241 ASSYSKNGFLYNRIGACTLVAADETVDRAFSQMKAAIRANYSNPPAHGASVATILSN 300
Db 243 SSFSKSFSLYGERVGCALSIVTESRDESARVLSQVKRVIRTNYSNPPTHGASVVSVLNS 302
Qy 301 DALRAIWQELTDMRORTORMQOLFVNTLQEKGANRDFSPILKQNGMFSFGLTKEOVLR 360
Db 303 PELRALWEQELGEMDRDRIDMLAMVEQLAAGHAKKEDFSFVGRQGRGMFSGLITADQVER 362
Qy 361 LREBFGVYVASGRVNVAGMTDPDNNAPLCEAIVAVL 396
Db 363 LKTEFGIYAVSTGRICVAALNKSNIETITKAIVQVL 398

RESULT 11
D83535
A: aromatic amino acid aminotransferase PA0870 [imported] - Pseudomonas aeruginosa (strain
C: Species: Pseudomonas aeruginosa
C: Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C: Accession: D83535
R: Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.: Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A: Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A: Reference number: AB2950; MUID:20437337; PMID:10984043
A: Accession: D83535
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-399 <STO>
A: Cross-references: UNIPROT:P43336; UNIPARC:UPI000013195D; GB:AE004522; GB:AE004091; NIL
C: Experimental source: strain PA01
C: Genetics:
A: Gene: phbC; PA0870
C: Superfamily: aspartate aminotransferase

```

Query Match	44.5%;	Score	910.5;	DB	2;	Length	399;
Best Local Similarity	47.0%;	Pred. NO.	1.1e-63;				
Matches	185;	Conservative	62;	Mismatches	146;	Indels	1;
							Gaps 1;
Qy	2	FENITAPADPILGLADLFRADERGKINLGTGVYKDETKCTPVLTSVKKAEQYLLLENET	61				
Db	4	FAKVARVPGDPILGLLDAYRNDPRADKDLGVGVYKDAQGLTTPILRSVKLAERQRLVEOET	63				
Qy	62	TKNYLGTIDGIPFGRCRTOELLFGKGSALINDRKRTAQTQPGGTGALRVAAADF LAKNTSVK	121				
Db	64	TKSVGVGGHDALFAARLAELGALGAASPLLEQRADATQTPGGTGALRLAGDFIAHCLPGR	123				
Qy	122	RVWYSNFSWPNHKS VFNFSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGCC	181				
Db	124	GIWLSDDPTWPIHETLFAAAGLKVSHVYVSADNR-LDVEAMLGLERIPQGDVVLLHACC	182				
Qy	182	HNPTGIDPTLEQWQTLAQLSVEKGMPLDFDPAYQGFARGLEDAEGLRAFAAMHKELIVA	241				
Db	183	HNPTGFDLSHDDWRVLDVWRVRRELPLTDFAYQGFGDGLEEDAWVRFLFAGELPEVLVT	242				
Qy	242	SSYSKFNGLYNERVGACCTLVAADSTVDVRAFCSMKAAIRANYSNPPAHGASVWATILISND	301				
Db	243	SSCSKFNGLYDRVCGALLVCQAQNAKLTDLRSQLAF LARNLWSTTPPAHGAEVVAAILGDS	302				
Qy	302	ALRAIWEQELDMRQRIQRMROLFVNTLQEKGNARDFSFIIKQNGMFSFSGLTKEQVURL	361				

Db 303 ELKGLWQBEVEGMRISRLASRLGLVEALPHGLAERFAHVGQAQRGMFSYTGLSPOQVARL 362
Qy 362 REEFQGVYAVASGRVNVAGTDPDNMAPLCEAIVAV 395
Db 363 RDEHAVLVSSGRANVAGLDARRDLRLAQAQV 396
RESULT 12
C82855
aromatic-amino-acid transaminase (EC 2.6.1.57) tyrB XF0036 [similarity] - Xylella fastidiosa
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: C82855
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: C82855
A:Status: preliminary
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <SIN>
A:Cross-references: UNIPROT:Q9PHA8; UNIPARC:UPI00000C229A; GB:AE003858; GB:AE003849; NID
A:Experimental source: strain 985C
R:Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Kramae, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matuskuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshakano, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0036
C:Superfamily: aspartate aminotransferase
C:Keywords: aminotransferase
Query Match 44.4%; Score 909; DB 2; Length 400;
Best Local Similarity 45.7%; Pred. No. 1.4e-63;
Matches 181; Conservative 76; Mismatches 137; Indels 2; Gaps 2;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINIGIGYKDETGTPTVLTSVKKAQYLLNE 60
Db 3 LFTDVELVGPDPILSLNDYNTADTRTNKVNIGIGYCDSEGCIPLLRAVQQVEQLAKHP 62
Qy 61 TTKVYLGIDGTEFCRCCTOELLFGKGSALINDKRAKTATCGGTGALRVAADFLAKNTSV 120
Db 63 KPRGYLPIDGLPAYIKATQQLFLFGVDSPLLTAGRVATTSQTITGGSGALRVTAEVLLKQVLP 122
Qy 121 KRVVYNSPNWPNHKSVPNSAGLEVREYAYDAENHTLDFDALINSNEAQAGDVVLFHGC 180
Db 123 ATVALSRFSWENHRALFTAAEFKEDITYFTDLGHAVDFTGMVADLAKLPKPTVVLLHGC 182
Qy 181 CHNPTGIDPTLEQWOTLAQLSVKGLWPLDFPAYQGFARGLEEDAEGLRAFAAMH-KELI 239
Db 183 CHNPTGADLRDQWQLVALFQERQLLPCIDLAYQFNGQIDADAYAIRLLAEGISNVY 242
Qy 240 VASSYSKNFGLYNERVGVACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILS 299
Db 243 VANSYSKFSYGERVGGSLIVASNTEQAQISQVKRIIRTIYSSPSAHGAYILVAGVLN 302
Qy 300 NDALRAIWEQELTDMRQIRMRQLFVNTLOEKGANRDFSFIIKQNGMFSFSGLTKEQVL 359
Db 303 SHELTLWEQELTQWREIHGLRAGLVARLLTALGA-PEFDFIQRAQNGFSYSGLSKIQVD 361
Qy 360 RLREBEFGVYAVASGRVNVAGTDPDNMAPLCEAIVAV 395

Db 362 RLREBEFGIYVSSGRICVAALSQHKLEYVAQAQVVKV 397
RESULT 13
JC5124
aspartate transaminase (EC 2.6.1.1), cytosolic - rice
N:Alternate names: cytoplasmic aspartate aminotransferase; cytoplasmic glutamic-oxaloac
C:Species: Oryza sativa (rice)
C>Date: 02-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: JC5124
R:Song, J.; Yamamoto, K.; Shomura, A.; Yano, M.; Minobe, Y.; Sasaki, T.
DNA Res. 3, 303-310, 1996
A:Title: Characterization and mapping of cDNA encoding aspartate aminotransferase in ric
A:Reference number: JC5124; MUID:97191542; PMID:9039500
A:Accession: JC5124
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-407 <SON>
A:Cross-references: UNIPROT:P37833; UNIPARC:UPI0000001116; GB:D14673; NID:g287297; PIDN:
C:Comment: This enzyme catalyzes the reversible transamination between a dicarboxylic ac
C:Genetics:
A:Map position: 1
C:Superfamily: aspartate aminotransferase
C:Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate
F:253/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
Query Match 44.4%; Score 909; DB 2; Length 407;
Best Local Similarity 45.0%; Pred. No. 1.5e-63;
Matches 179; Conservative 69; Mismatches 148; Indels 2; Gaps 2;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINIGIGYKDETGTPTVLTSVKKAQYLLNE 60
Db 6 VFAGLAQAPEDPILGVTVAYNKDPSPVKVNLGVGAYRTEEGKPLVLNVVRAEQMLINNP 65
Qy 61 T-TKNVYLGIDGTEFCRCCTOELLFGKGSALINDKRAKTATCGGTGALRVAADFLAKNTS 119
Db 66 SRVKEYLPITGLADFNLKSLAKLIFGADSPAIQENRVATVOCLSGTSGLRVGGEFLLARHYH 125
Qy 120 KRVVYNSPNWPNHKSVPNSAGLEVREYAYDAENHTLDFDALINSNEAQAGDVVLFHGC 179
Db 126 ERTIYIPQWTGNHPKVFTLAGLTVRSYRYYPATRGDLDFQGLLEDLGSAPSGAIVLLHA 185
Qy 180 CHNPTGIDPTLEQWOTLAQLSVKGLWPLDFPAYQGFARG-LEEDAEGLRAFAAMH-KEL 238
Db 186 CAHNPTGVDPDLTQWEQIRQLMRSKALLPFFDSAYQGFASGSLDQDAQSRMVFADGGEL 245
Qy 239 IVASSYSKNFGLYNERVGVACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATIL 298
Db 246 LMAQSYAKNMGLYGERVGLSIVCGSADVAVRVESQKLVIRPMYSNPPIHGASIVATIL 305
Qy 299 SNDALRAIWEQELTDMRQIRMRQLFVNTLOEKGANRDFSFIIKQNGMFSFSGLTKEQV 358
Db 306 KDSAMFNWETVELKGMADRIISMRQLFDALKTRTETPGDWSHIILKQIGMFTTGLNSDQV 365
Qy 359 RLREBEFGVYAVASGRVNVAGTDPDNMAPLCEAIVAVL 396
Db 366 AFMRQEVHYWTS DGRISMAGLSGRRTIPHLADAIHAAV 403
RESULT 14
S53303
aspartate transaminase (EC 2.6.1.1) AAT2 - proso millet
N:Alternate names: aspartate aminotransferase, cytosolic
C:Species: Panicum miliaceum (proso millet)
C>Date: 14-Jul-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S53303; S22378; S40067; S18891
R:Taniguchi, M.; Mori, J.; Sugiyama, T.
Plant Mol. Biol. 26, 723-734, 1994
A:Title: Structure of genes that encode isozymes of aspartate aminotransferase in Panicu
A:Reference number: S53303; MUID:95036052; PMID:7948926
A:Accession: S53303
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-409 <TAN>
A:Cross-references: UNIPROT:Q42391; UNIPARC:UPI00000AA577; EMBL:D25322; NID:G435456; PID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
R:Taniguchi, M.; Sawaki, H.; Sasakawa, H.; Hase, T.; Sugiyama, T.
Eur. J. Biochem. 204, 611-620, 1992
A:Title: Cloning and sequence analysis of cDNA encoding aspartate aminotransferase isozy
A:Reference number: S22377; MUID:92174917; PMID:1541276
A:Accession: S22378
A:Molecule type: mRNA
A:Residues: 1-409 <TAN>
A:Cross-references: UNIPARC:UPI00000AA577; EMBL:X63429; NID:G20598; PIDN:CAA45023.1; PID
A:Accession: S40067
A:Molecule type: protein
A:Residues: 72-84; 'X', 130-136; 213-230 <TAN2>
A:Cross-references: UNIPARC:UPI00001754F5; UNIPARC:UPI00001754F6; UNIPARC:UPI00001754F7
C:Genetics:
A:Gene: AAT2
A:Introns: 22/3; 46/3; 67/2; 95/2; 128/3; 169/1; 223/3; 269/3; 310/2; 342/1; 382/2
C:Superfamily: aspartate aminotransferase
C:Keywords: aminotransferase; cytosol; phosphoprotein; pyridoxal phosphate
F:255/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted
Query Match 44.4%; Score 908; DB 2; Length 409;
Best Local Similarity 45.5%; Pred. No. 1.8e-63;
Matches 181; Conservative 67; Mismatches 148; Indels 2; Gaps 2;
QY 1 MFENITAAPADPILGLADLFRADERPGKINIGVYKDETGTPTVLTSVKKAQVLLNE 60
DB 8 VFAGIAQAPEDPILGVTVAFNKDPSVKINLGVGAYRTTEEGKPLVLNVVRAEQMLINDP 67
QY 61 T-TKNYLGIDGPEFCRGTQELLFGKGSALINDKARTATQTPGTCALRVAADFLAKNTS 119
DB 68 SRVKEYLPITGLAEYNKLSAKLIFGADSPAIOENRVATVQCLSGTSLRVGGEFLAKHYH 127
QY 120 VKRVMVSNPSPNHSKVSFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAQGVVLFHG 179
DB 128 ERTIVIPVTWGNHPKVFTLAGLTVRSYRYYPATRGDLDFNGLLDLSAPLGSIVLLHA 187
QY 180 CCHNPTGIDPTLEOMOTLAQLSVEKGLPLDFPAYOGFARG-LEEDAELRAFAAMHKL 238
DB 188 CAHNPTGVDPDTIQEQIRQLMRSKSLPFFDPSAYOGFASGLDADAQSVRIFFVADGGEL 247
QY 239 IVASSYSKNFGLYNERVGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVVATIL 298
DB 248 LAAQSYAKNMGYGERVGSALVCSADVAARVESQKLVIRPMYSSPLHGSIVVATIL 307
QY 299 SNDALRAIWEQELTDMRORIORMQLFVNTLOEKGANRDFPIIKONGMFSFSLTKQOV 358
DB 308 KDSMFHEWTVELKAMADRIISMROQLFDALSRGTGPDWSHIIKQIGMFTFTGLNSEQV 367
QY 359 LRLREFGVYAVASGRVNVAGMTDPNMAPLCEAIVAVL 396
DB 368 APMROEYHIYMTSDGRISMAGLNKNVPHLADAIHAAV 405

RESULT 15
T14311
aspartate transaminase (EC 2.6.1.1), cytosolic [similarity] - carrot
N:Alternate names: aspartate aminotransferase
C:Species: Daucus carota (carrot)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T14311
R:Turano, F.J.; Weisemann, J.M.; Matthews, B.F.
submitted to the EMBL Data Library, July 1992
A:Description: Identification and expression of a cDNA clone encoding aspartate aminotra
A:Reference number: Z17971
A:Accession: T14311
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-405 <TNR>
A:Cross-references: UNIPROT:P28734; UNIPARC:UPI00001250D4; EMBL:M92660; NID:G167545; PID
C:Function:

A:Description: reversibly catalyzes transamination between L-aspartate and 2-oxoglutarate
C:Superfamily: aspartate aminotransferase
C:Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate
F:251/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted
Query Match 44.2%; Score 904; DB 2; Length 405;
Best Local Similarity 44.7%; Pred. No. 3.6e-63;
Matches 178; Conservative 67; Mismatches 151; Indels 2; Gaps 2;
QY 1 MFENITAAPADPILGLADLFRADERPGKINIGVYKDETGTPTVLTSVKKAQVLLNE 60
DB 4 VFANVVRAPEDPILGVTVAYHKDQSPKINLGVGAYRTTEEGKPLVLNVVKAQVLLNQ 63
QY 61 T-TKNYLGIDGPEFCRGTQELLFGKGSALINDKARTATQTPGTCALRVAADFLAKNTS 119
DB 64 SRVKEYLPITGLADLDFNKLAKLIFGADSPAIOENRVATVQCLSGTSLRVGGEFLARHYH 123
QY 120 VKRVMVSNPSPNHSKVSFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAQGVVLFHG 179
DB 124 EHTVYIPQPTWGNHPKIFTLAGLSVKTYRYYPATRGDLDFEGMLDLSAPLGAIVLLHA 183
QY 180 CCHNPTGIDPTLEOMOTLAQLSVEKGLPLDFPAYOGFARG-LEEDAELRAFAAMHKL 238
DB 184 CAHNPTGVDPDTIQEQIRQLMRSKSLPFFDPSAYOGFASGLDADAQSVRIFFVADGGEC 243
QY 239 IVASSYSKNFGLYNERVGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVVATIL 298
DB 244 LAAQSYAKNMGYGERVGSALVCSADVAARVESQKLVIRPMYSSPLHGSIVVATIL 303
QY 299 SNDALRAIWEQELTDMRORIORMQLFVNTLOEKGANRDFPIIKONGMFSFSLTKQOV 358
DB 304 KDGDLYNWTLLEKAMADRIISMROELFNALQAKGTGPDWSHIVKQIGMFTFTGLNSEQV 363
QY 359 LRLREFGVYAVASGRVNVAGMTDPNMAPLCEAIVAVL 396
DB 364 TFMTEYHIYMTSDGRISMAGLSRTVPHLADAIHAAV 401

Search completed: March 14, 2006, 18:00:45
Job time : 41 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 17:53:05 ; Search time 230 seconds
(without alignments)
1214.736 Million cell updates/sec

Title: US-10-673-786A-2
Perfect score: 2045
Sequence: 1 MFENITAPADPILGLDLF.....VAGMTPDNWPLCEAIVAVL 396

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2045	100.0	396	1	AAT ECOLI
2	2037	99.6	396	2	Q8XDF3 ECOLI
3	2033	99.4	396	2	Q8FJ99 ECOLI
4	2026	99.1	396	2	Q83LNA SHIFL
5	1982	96.9	396	1	AAT_SALTY
6	1979	96.8	396	2	Q57R00 SALCH
7	1978	96.7	396	1	AAT_SALTY
8	1978	96.7	396	2	Q5PGE8 SALPA
9	1780	87.0	396	2	Q6D451 ERWCT
10	1775	86.8	396	2	Q8ZG95 YERPE
11	1775	86.8	396	2	Q6CG99 YERPE
12	1775	86.8	401	2	Q8D050 YERPE
13	1682	82.2	396	2	Q7N625 PHOLL
14	1349	66.0	396	2	Q5D191 9PAST
15	1343	65.7	396	2	Q4QL54 HAEPHIL
16	1336	65.3	399	2	Q5B4R3 VIBF1
17	1333	65.2	396	2	Q6LPQ0 PHOFR
18	1329	65.0	396	1	AAT_HAEIN
19	1325	64.8	397	2	Q8EEM8 SHON
20	1325	64.8	414	2	Q87NG8 VIBPA
21	1324	64.7	396	2	Q8RNA0 AERHY
22	1318	64.4	399	2	Q8DAF6 VIBVU
23	1317	64.4	396	2	Q5F6U3 NEIGI
24	1316	64.4	399	2	Q7MJR4 VIBVU
25	1315	64.3	413	2	Q9KSG3 VIBCH
26	1311	64.1	397	2	Q9JVS3 NEIMA
27	1310	64.1	397	2	Q9KOP5 NEIMB
28	1299	63.5	396	2	Q9CN25 PASMU
29	1279	62.5	420	2	Q65TS2 MANHM
30	1257	61.5	396	2	Q7VLL0 HAEDU
31	1125.5	55.0	406	2	Q7VR08 CANBF

32	1081	52.9	397	2	Q7UHY8 RHOB
33	1055	51.6	398	2	Q6AQM7 DESPS
34	1051.5	51.4	397	2	Q53137 MORSP
35	1044	51.1	196	2	Q5INQ0 ECOLI
36	1044	51.1	196	2	Q5INQ8 ECOLI
37	1044	51.1	196	2	Q5INQ8 SHIDY
38	1040	50.9	196	2	Q5INT6 SHIBO
39	1039	50.8	195	2	Q5INR1 ECOLI
40	1039	50.8	195	2	Q5INQ8 SHIDY
41	1038	50.8	195	2	Q5INS6 SHIDY
42	1038	50.8	195	2	Q5INT8 SHISO
43	1038	50.8	195	2	Q5INT9 SHIBO
44	1033	50.5	194	2	Q5INH8 ECOLI
45	1033	50.5	194	2	Q5INS5 SHIDY

ALIGNMENTS

RESULT 1
AAT ECOLI
ID AAT ECOLI STANDARD; PRT; 396 AA.
AC P00509;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Aspartate aminotransferase (EC 2.6.1.1) (Transaminase A) (ASPART).
GN Name=aspC; OrderedLocusNames=b0928;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=85289110; PubMed=3897210;
RA Kuramitsu S., Okuno S., Ogawa T., Ogawa H., Kagamiyama H.;
RT "Aspartate aminotransferase of Escherichia coli: nucleotide sequence of the aspC gene."
RL J. Biochem. 97:1259-1262(1985).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=86242111; PubMed=3521591;
RA Fotheringham I.G., Dacey S.A., Taylor P.P., Smith T.J., Hunter M.G., Finlay M.E., Primrose S.B., Parker D.M., Edwards R.M.;
RT "The cloning and sequence analysis of the aspC and tyxB genes from Escherichia coli K12. Comparison of the primary structures of the aspartate aminotransferase and aromatic aminotransferase of E. coli with those of the pig aspartate aminotransferase isoenzymes."
RL Biochem. J. 234:593-604(1986).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A., Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K., Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N., Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 12.7-28.0 min region on the linkage map."
RL DNA Res. 3:137-155(1996).
RN [5]

```

RP RX MEDLINE=84256832; PubMed=6378205;
RA Kondo K., Wakabayashi S., Yagi T., Kagamiyama H.;
RT "The complete amino acid sequence of aspartate aminotransferase from
RL Escherichia coli: sequence comparison with pig isoenzymes.";
RN Biochem. Biophys. Res. Commun. 122:62-67(1984).
RP PROTEIN SEQUENCE.
RX MEDLINE=87250482; PubMed=3298240;
RA Kondo K., Wakabayashi S., Kagamiyama H.;
RT "Structural studies on aspartate aminotransferase from Escherichia
RL coli. Covalent structure.";
RN J. Biol. Chem. 262:8648-8659(1987).
RP PROTEIN SEQUENCE OF 1-12.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RL in the genome of Escherichia coli K-12.";
RN Electrophoresis 18:1259-1313(1997).
RP MUTAGENESIS OF TYR-65.
RX MEDLINE=91329346; PubMed=1868057;
RA Inoue K., Kuramitsu S., Okamoto A., Hirotsu K., Higuchi T.,
RA Kagamiyama H.;
RT "Site-directed mutagenesis of Escherichia coli aspartate
RL aminotransferase: role of Tyr70 in the catalytic processes.";
RN Biochemistry 30:7796-7801(1991).
RP MUTAGENESIS OF HIS-133.
RX MEDLINE=91177849; PubMed=2007566;
RA Yano T., Kuramitsu S., Tanase S., Morino Y., Hiromi K., Kagamiyama H.;
RT "The role of His143 in the catalytic mechanism of Escherichia coli
RL aspartate aminotransferase.";
RN J. Biol. Chem. 266:6079-6085(1991).
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF MUTANT ALA-246.
RX MEDLINE=90105323; PubMed=2513875;
RA Smith D.L., Almo S.C., Toney M.D., Ringe D.;
RT "2.8-A-resolution crystal structure of an active-site mutant of
RL aspartate aminotransferase from Escherichia coli.";
RN Biochemistry 28:8161-8177(1989).
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS), AND MUTAGENESIS OF ARG-374.
RX MEDLINE=91129283; PubMed=1593208;
RA Danishefsky A.T., Omufu J.J., Petsko G.A., Ringe D.;
RT "Activity and structure of the active-site mutants R386Y and R386F of
RL Escherichia coli aspartate aminotransferase.";
RN Biochemistry 30:1980-1985(1991).
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF MUTANT.
RX MEDLINE=99107891; PubMed=9891001; DOI=10.1074/jbc.274.4.2344;
RA Oue S., Okamoto A., Yano T., Kagamiyama H.;
RT "Redesigning the substrate specificity of an enzyme by cumulative
RL effects of the mutations of non-active site residues.";
RN J. Biol. Chem. 274:2344-2349(1999).
CC -!- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
CC L-glutamate.
CC -!- COFACTOR: Pyridoxal phosphate.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the class-I pyridoxal-phosphate-dependent
CC aminotransferase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; X03629; CAA2279.1; -; Genomic DNA.
CC EMBL; X05904; CAA29333.1; -; Genomic DNA.
DR EMBL; U00096; AAC74014.1; -; Genomic DNA.
DR EMBL; D90730; BAA35674.1; -; Genomic DNA.
DR EMBL; D90731; BAA35680.1; -; Genomic DNA.
DR PIR; A00598; XNRCDD.
DR PDB; 1AAM; X-ray; @=1-396.
DR PDB; 1AAW; X-ray; @=1-396.
DR PDB; 1AHE; X-ray; A/B=1-396.
DR PDB; 1AHF; X-ray; A/B=1-396.
DR PDB; 1AHG; X-ray; A/B=1-396.
DR PDB; 1AHX; X-ray; A/B=1-396.
DR PDB; 1AHY; X-ray; A/B=1-396.
DR PDB; 1AIA; X-ray; A/B=1-396.
DR PDB; 1AIB; X-ray; A/B=1-396.
DR PDB; 1AIC; X-ray; A/B=1-396.
DR PDB; 1AMQ; X-ray; @=1-396.
DR PDB; 1AMR; X-ray; @=1-396.
DR PDB; 1AMS; X-ray; @=1-396.
DR PDB; 1ARG; X-ray; A/B=1-396.
DR PDB; 1ARH; X-ray; A/B=1-396.
DR PDB; 1ARI; X-ray; A/B=1-396.
DR PDB; 1ARS; X-ray; @=1-396.
DR PDB; 1ART; X-ray; @=1-396.
DR PDB; 1ASA; X-ray; @=1-396.
DR PDB; 1ASB; X-ray; @=1-396.
DR PDB; 1ASC; X-ray; @=1-396.
DR PDB; 1ASD; X-ray; @=1-396.
DR PDB; 1ASE; X-ray; @=1-396.
DR PDB; 1ASF; X-ray; @=1-396.
DR PDB; 1ASG; X-ray; @=1-396.
DR PDB; 1ASL; X-ray; A/B=1-396.
DR PDB; 1ASM; X-ray; A/B=1-396.
DR PDB; 1ASN; X-ray; A/B=1-396.
DR PDB; 1B4X; X-ray; A=1-396.
DR PDB; 1BOA; X-ray; A/B=1-396.
DR PDB; 1BOD; X-ray; A/B=1-396.
DR PDB; 1C9C; X-ray; A=1-396.
DR PDB; 1CQ6; X-ray; A=1-396.
DR PDB; 1CQ7; X-ray; A=1-396.
DR PDB; 1CQ8; X-ray; A=1-396.
DR PDB; 1CZC; X-ray; A=1-396.
DR PDB; 1CZE; X-ray; A=1-396.
DR PDB; 1G4V; X-ray; A=1-396.
DR PDB; 1G4X; X-ray; A=1-396.
DR PDB; 1G7W; X-ray; A=1-396.
DR PDB; 1G7X; X-ray; A=1-396.
DR PDB; 1IX6; X-ray; A=1-396.
DR PDB; 1IX7; X-ray; A=1-396.
DR PDB; 1IX8; X-ray; A=1-396.
DR PDB; 1Q1R; X-ray; A=1-396.
DR PDB; 1Q1T; X-ray; A=1-396.
DR PDB; 1SPA; X-ray; @=1-396.
DR PDB; 1TOE; X-ray; A=1-396.
DR PDB; 1TOG; X-ray; A/B=1-396.
DR PDB; 1TOI; X-ray; A=1-396.
DR PDB; 1TOJ; X-ray; A=1-396.
DR PDB; 1TOK; X-ray; A/B=1-396.
DR PDB; 1YOO; X-ray; @=1-396.
DR PDB; 2AAT; X-ray; @=1-396.
DR PDB; 3AAT; X-ray; @=1-396.
DR PDB; 5EAA; X-ray; A=1-396.
DR SWISS-2DPAGE; P00509; COLI.
DR EC02DBASE; F039.6; 6TH EDITION.
DR EC02DBASE; F039.7; 6TH EDITION.
DR EchoBASE; EB0094; -.
DR Ecogene; EG10096; aspc.
DR InterPro; IPR004839; Aminotrans_I/II.
DR InterPro; IPR000796; Asp.trans.
DR InterPro; IPR004838; Nhttransf_1_BS.
DR PANTHER; PTHR11879; Asp.trans; 1.
DR Pfam; PF00155; Aminotran_1_2; 1.
DR PRINTS; PR00799; TRANSAMINASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.

```


KW 3D-structure; Amino transferase; Complete proteome;
 KW Direct protein sequencing; Pyridoxal phosphate; Transferase.

Query Match 100.0%; Score 2045; DB 1; Length 396;
 Best Local Similarity 100.0%; Pred. No. 3e-146;
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFENITAAPADPILGLADLFRADERPGKINLGIVGYKDETGTPTVLTSVKKAEOYLLENE 60
 DB 1 MFENITAAPADPILGLADLFRADERPGKINLGIVGYKDETGTPTVLTSVKKAEOYLLENE 60
 QY 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAARTAOPTPGTGALRVAADFLAKNTSV 120
 DB 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAARTAOPTPGTGALRVAADFLAKNTSV 120
 QY 121 KRVMVSNPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLEAQAQDVVLFHGC 180
 DB 121 KRVMVSNPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLEAQAQDVVLFHGC 180
 QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGLPLFDFAVQGFARGLEEDAEGLRAFAAMHKELIV 240
 DB 181 CHNPTGIDPTLEQWOTLAQLSVEKGLPLFDFAVQGFARGLEEDAEGLRAFAAMHKELIV 240
 QY 241 ASSYKNGFLYNERVAGCTLVADSETVDRAFSONKAAIRANYSNPPAHGASVVAATILSN 300
 DB 241 ASSYKNGFLYNERVAGCTLVADSETVDRAFSONKAAIRANYSNPPAHGASVVAATILSN 300
 QY 301 DALRAIWEQELTDMQRQIRMQRLFVNTLOEGKANRDFSFIKONGMFSGLTKEQVLR 360
 DB 301 DALRAIWEQELTDMQRQIRMQRLFVNTLOEGKANRDFSFIKONGMFSGLTKEQVLR 360
 QY 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
 DB 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 2

QXDF3 ECO57
 ID Q8XDF3 ECO57 PRELIMINARY; PRT; 396 AA.
 AC Q8XDF3_07AG50;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
 DE Aspartate aminotransferase.
 GN Name=aspC; OrderedLocusNames=ECs1011, z1275;
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.P., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postfai J., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
 RL Nature 409:529-533 (2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=O157:H7 / Sakai / RIMD 050952 / EHEC;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogatawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12."
 RL DNA Res. 8:11-22 (2001).
 RL EMBL; AE005174; AAG55413.1; -; Genomic_DNA.

DR EMBL; BA000007; BAB34434.1; -; Genomic_DNA.
 DR PIR; A85619; A85619.
 DR PIR; C90755; C90755.
 DR HSSP; P00509; 1ART.
 DR SMR; Q8XDF3; 1-396.
 DR GO; GO:0008483; Fibrinase activity; IEA.
 DR GO; GO:0006520; Pi amino acid metabolism; IEA.
 DR GO; GO:0009058; Biosynthesis; IEA.
 DR InterPro; IPR004839; Amino transferase; IEA.
 DR InterPro; IPR000796; Asp trans I/II.
 DR InterPro; IPR004838; NHz trans I_BS.
 DR Pfam; PF00155; Amino transferase; IEA.
 DR PRINTS; PR00799; TRANSAMINASE.
 DR PROSITE; PS00105; AA TRANSFER CLASS 1; 1.
 KW Amino transferase; Complete proteome; Transferase.
 SQ SEQUENCE 396 AA; 43645 MW; 81C2063CE24DF08 CRC64;

Query Match 99.6%; Score 2037; DB 2; Length 396;
 Best Local Similarity 99.5%; Pred. No. 1.2e-145;
 Matches 394; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFENITAAPADPILGLADLFRADERPGKINLGIVGYKDETGTPTVLTSVKKAEOYLLENE 60
 DB 1 MFENITAAPADPILGLADLFRADERPGKINLGIVGYKDETGTPTVLTSVKKAEOYLLENE 60
 QY 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAARTAOPTPGTGALRVAADFLAKNTSV 120
 DB 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAARTAOPTPGTGALRVAADFLAKNTSV 120
 QY 121 KRVMVSNPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLEAQAQDVVLFHGC 180
 DB 121 KRVMVSNPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLEAQAQDVVLFHGC 180
 QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGLPLFDFAVQGFARGLEEDAEGLRAFAAMHKELIV 240
 DB 181 CHNPTGIDPTLEQWOTLAQLSVEKGLPLFDFAVQGFARGLEEDAEGLRAFAAMHKELIV 240
 QY 241 ASSYKNGFLYNERVAGCTLVADSETVDRAFSONKAAIRANYSNPPAHGASVVAATILSN 300
 DB 241 ASSYKNGFLYNERVAGCTLVADSETVDRAFSONKAAIRANYSNPPAHGASVVAATILSN 300
 QY 301 DALRAIWEQELTDMQRQIRMQRLFVNTLOEGKANRDFSFIKONGMFSGLTKEQVLR 360
 DB 301 DALRAIWEQELTDMQRQIRMQRLFVNTLOEGKANRDFSFIKONGMFSGLTKEQVLR 360
 QY 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
 DB 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 3

Q8FJ99 ECO6
 ID Q8FJ99 ECO6 PRELIMINARY; PRT; 396 AA.
 AC Q8FJ99;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Aspartate aminotransferase (EC 2.6.1.1).
 GN Name=aspC; OrderedLocusNames=c1070;
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=217992;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
 RX MEDLINE=2338234; PubMed=12471157; DOI=10.1073/pnas.252529799;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesech P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.P., Rose D.J., Zhou S., Schwartz F.R.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic Escherichia coli.";

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AE016758; AAN79538.1; -; Genomic_DNA.
DR HSSP; P00509; IART.
DR SMR; Q8FJ99; 1-396.
DR GO; GO:0004069; P:aspartate transaminase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR004839; AminoTrans_I/II.
DR InterPro; IPR00796; AspTrans.
DR InterPro; IPR004838; NHTransf_1_BS.
DR Pfam; PF00155; AminoTrans_1_2_1.
DR PRINTS; PR00799; TRANSAMINASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
DR Complete proteome.
KW SEQUENCE 396 AA; 43568 MW; 9FA46ECB413FC679 CRC64;

Query Match 99.4%; Score 2033; DB 2; Length 396;
Best Local Similarity 99.0%; Pred. No. 2.4e-145;
Matches 392; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINLGIGYKDETKPTVLTSVKKAQVLLNE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINLGIGYKDETKPTVLTSVKKAQVLLNE 60

Qy 61 TTKNYLGIDGPEFCRCQTELLFGKGSALINDKARTAQTPGGTGALRVAADFLAKNTSV 120
Db 61 TTKNYLGIDGLPEFCRCQTELLFGKGSALINDKARTAQTPGGTGALRVAADFLAKNTSV 120

Qy 121 KRVMVSNPSPNKHVSFNSAGLEVREYAYDAENHTLDFDALINSLEAQAQGVVLPFGC 180
Db 121 KRVMVSNPSPNKHVSFNSAGLEVREYAYDAENHTLDFDALINSLEAQAQGVVLPFGC 180

Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKGLPLDFAYQGFARGLEEDAEGLRFAAAMHKEIV 240
Db 181 CHNPTGIDPTLEQWQTLAQLSVEKGLPLDFAYQGFARGLEEDAEGLRFAAAMHKEIV 240

Qy 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAIRANYSNPPAHGASVVATILSN 300

Qy 301 DALRAIWEQELTDMRQIRMRQLFVNTLQEKGNRDFSLIKONGMFSFGLTKQVLR 360
Db 301 DALRAIWEQELTDMRQIRMRQLFVNTLQEKGNRDFSLIKONGMFSFGLTKQVLR 360

Qy 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 4
ID Q83LN4 SHIFL PRELIMINARY; PRT; 396 AA.
AC Q83LN4_07C286;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Aspartate aminotransferase.
GN Name=aspC; OrderedLocusNames=S0989; SF0925;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=2227406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
```

```

RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/IAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AE005674; AAN42554.1; -; Genomic_DNA.
DR EMBL; AE016981; AAP16440.1; -; Genomic_DNA.
DR HSSP; P00509; IART.
DR SMR; Q83LN4; 1-396.
DR GO; GO:0008483; P:transaminase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR004839; AminoTrans_I/II.
DR InterPro; IPR00796; AspTrans.
DR InterPro; IPR004838; NHTransf_1_BS.
DR Pfam; PF00155; AminoTrans_1_2_1.
DR PRINTS; PR00799; TRANSAMINASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
DR AminoTransferase; Complete proteome; Transferase.
KW SEQUENCE 396 AA; 43617 MW; 621C1FB9FBA4DD5A CRC64;

Query Match 99.1%; Score 2026; DB 2; Length 396;
Best Local Similarity 99.2%; Pred. No. 8.2e-145;
Matches 393; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINLGIGYKDETKPTVLTSVKKAQVLLNE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINLGIGYKDETKPTVLTSVKKAQVLLNE 60

Qy 61 TTKNYLGIDGPEFCRCQTELLFGKGSALINDKARTAQTPGGTGALRVAADFLAKNTSV 120
Db 61 TTKNYLGIDGLEFCRCQTELLFGKGSALINDKARTAQTPGGTGALRVAADFLAKNTSV 120

Qy 121 KRVMVSNPSPNKHVSFNSAGLEVREYAYDAENHTLDFDALINSLEAQAQGVVLPFGC 180
Db 121 KRVMVSNPSPNKHVSFNSAGLEVREYAYDAENHTLDFDALINSLEAQAQGVVLPFGC 180

Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKGLPLDFAYQGFARGLEEDAEGLRFAAAMHKEIV 240
Db 181 CHNPTGIDPTLEQWQTLAQLSVEKGLPLDFAYQGFARGLEEDAEGLRFAAAMHKEIV 240

Qy 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAIRANYSNPPAHGASVVATILSN 300

Qy 301 DALRAIWEQELTDMRQIRMRQLFVNTLQEKGNRDFSLIKONGMFSFGLTKQVLR 360
Db 301 DALRAIWEQELTDMRQIRMRQLFVNTLQEKGNRDFSLIKONGMFSFGLTKQVLR 360

Qy 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 5
ID AAT_SALTY STANDARD; PRT; 396 AA.
AC P58661;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Aspartate aminotransferase (EC 2.6.1.1) (Transaminase A) (AspAT).
GN Name=aspC; OrderedLocusNames=STM0998;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
```

OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN (1)
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11877609; DOI=10.1038/35101614;
 RA McClelland M., Sanderson K.E., Speth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 LT2.";
 RL Nature 413:852-856 (2001).
 CC -!- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
 CC L-glutamate.
 CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to the class-I pyridoxal-phosphate-dependent
 CC aminotransferase family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AE008743; AAL19932.1; -; Genomic_DNA.
 DR HSSP; P00509; IART.
 DR SMR; P58661; 1-396.
 DR StyGene; SG77777; aspC.
 DR InterPro; IPR004839; Aminotrans_I/II.
 DR InterPro; IPR000796; Asp_trans.
 DR InterPro; IPR004838; Nhrtransf_BS.
 DR PANTHER; PTHR11879; Asp_trans; 1.
 DR Pfam; PF00155; Aminotran 1.2; 1.
 DR PRINTS; PR00799; TRANSAMINASE.
 DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
 DR Aminotransferase; Complete proteome; Pyridoxal phosphate; Transferase.
 KW ACT SITE 374 246
 FT BINDING 246 246 Pyridoxal phosphate (covalent) (By
 FT similarity).
 FT SEQUENCE 396 AA; 43521 MW; 960940D0148D02FB CRC64;
 SQ
 Query Match 96.9%; Score 1982; DB 1; Length 396;
 Best Local Similarity 96.0%; Pred. No. 1.8e-141;
 Matches 380; Conservative 11; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MFENITAAPADPILGLADLFRADRRPGKINLGIVGVDGTGTPVLTSTVKKAEQYLLEN 60
 DB 1 MFENITAAPADPILGLADLFRADRRPGKINLGIVGVDGTGTPVLTSTVKKAEQYLLEN 60
 QY 61 TTKNYLGIDGIEFGRCCTOELLFGKGSALINDKARTATPGGTGALRVAADFLAKNTSV 120
 DB 61 TTKNYLGIDGIEFGRCCTOELLFGKGSALINDKARTATPGGTGALRVAADFLAKNTSV 120
 QY 121 KRWVSNPSPNPKSVFNAGLEVEYAYDAENHTLDPDALINSNEAQAGDVVLFHGC 180
 DB 121 KRWVSNPSPNPKSVFNAGLEVEYAYDAENHTLDPDALINSNEAQAGDVVLFHGC 180
 QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFPAYOGFARGLEDEAGLFAFAHKLIV 240
 DB 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFPAYOGFARGLEDEAGLFAFAHKLIV 240
 QY 241 ASSYKNGFLYNERVAGCTLVAADETVDRAFQSKAAIRANYSNPPHAGSIVATILSN 300
 DB 241 ASSYKNGFLYNERVAGCTLVAADETVDRAFQSKAAIRANYSNPPHAGSIVATILSN 300
 QY 301 DALRAIWEQELTDMQRQIRMRQLFVNTLOEGKANRDFSFIKQNGMFSFGLTKQVLR 360
 DB 301 DALRAIWEQELTDMQRQIRMRQLFVNTLOEGKANRDFSFIKQNGMFSFGLTKQVLR 360
 QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFPAYOGFARGLEDEAGLFAFAHKLIV 240
 DB 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFPAYOGFARGLEDEAGLFAFAHKLIV 240
 QY 241 ASSYKNGFLYNERVAGCTLVAADETVDRAFQSKAAIRANYSNPPHAGSIVATILSN 300
 DB 241 ASSYKNGFLYNERVAGCTLVAADETVDRAFQSKAAIRANYSNPPHAGSIVATILSN 300
 QY 301 DALRAIWEQELTDMQRQIRMRQLFVNTLOEGKANRDFSFIKQNGMFSFGLTKQVLR 360
 DB 301 DALRAIWEQELTDMQRQIRMRQLFVNTLOEGKANRDFSFIKQNGMFSFGLTKQVLR 360

QY 361 LREEFGVYAVASGRVNVAGMTDPDNNAPLCEAIVAVL 396
 DB 361 LREEFGVYAVASGRVNVAGMTDPDNNAPLCEAIVAVL 396
 RESULT 6
 Q57R00_SALCH PRELIMINARY; PRT; 396 AA.
 AC Q57R00;
 DT 10-MAY-2005 (TREMBLrel. 30, Created)
 DT 10-MAY-2005 (TREMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
 DE Aspartate aminotransferase.
 GN Name=aspC; OrderedLocusNames=SC0955;
 OS *Salmonella cholerae-suis* (*Salmonella enterica*).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Salmonella*.
 OX NCBI_TaxID=591;
 RN (1)
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=SC-B67;
 RX PubMed=15781495;
 RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
 RA Wang H.-S., Lee Y.-S.;
 RT "The genome sequence of *Salmonella enterica* serovar *Choleraesuis*, a
 RT highly invasive and resistant zoonotic pathogen.";
 RL Nucleic Acids Res. 33:1690-1698 (2005).
 DR EMBL; AE017220; AAX64861.1; -; Genomic DNA.
 KW Aminotransferase; Complete proteome; Transferase.
 SQ SEQUENCE 396 AA; 43537 MW; 891640CF08B02FB CRC64;
 Query Match 96.8%; Score 1979; DB 2; Length 396;
 Best Local Similarity 95.7%; Pred. No. 3e-141;
 Matches 379; Conservative 12; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MFENITAAPADPILGLADLFRADRRPGKINLGIVGVDGTGTPVLTSTVKKAEQYLLEN 60
 DB 1 MFENITAAPADPILGLADLFRADRRPGKINLGIVGVDGTGTPVLTSTVKKAEQYLLEN 60
 QY 61 TTKNYLGIDGIEFGRCCTOELLFGKGSALINDKARTATPGGTGALRVAADFLAKNTSV 120
 DB 61 TTKNYLGIDGIEFGRCCTOELLFGKGSALINDKARTATPGGTGALRVAADFLAKNTSV 120
 QY 121 KRWVSNPSPNPKSVFNAGLEVEYAYDAENHTLDPDALINSNEAQAGDVVLFHGC 180
 DB 121 KRWVSNPSPNPKSVFNAGLEVEYAYDAENHTLDPDALINSNEAQAGDVVLFHGC 180
 QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFPAYOGFARGLEDEAGLFAFAHKLIV 240
 DB 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFPAYOGFARGLEDEAGLFAFAHKLIV 240
 QY 241 ASSYKNGFLYNERVAGCTLVAADETVDRAFQSKAAIRANYSNPPHAGSIVATILSN 300
 DB 241 ASSYKNGFLYNERVAGCTLVAADETVDRAFQSKAAIRANYSNPPHAGSIVATILSN 300
 QY 301 DALRAIWEQELTDMQRQIRMRQLFVNTLOEGKANRDFSFIKQNGMFSFGLTKQVLR 360
 DB 301 DALRAIWEQELTDMQRQIRMRQLFVNTLOEGKANRDFSFIKQNGMFSFGLTKQVLR 360
 QY 361 LREEFGVYAVASGRVNVAGMTDPDNNAPLCEAIVAVL 396
 DB 361 LREEFGVYAVASGRVNVAGMTDPDNNAPLCEAIVAVL 396
 RESULT 7
 AAT_SALTI STANDARD; PRT; 396 AA.
 AC Q56114;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Aspartate aminotransferase (EC 2.6.1.1) (Transaminase A) (AspAT).
 GN Name=aspC; OrderedLocusNames=STY1000, t1936;

OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=CT18;
 RA MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
 RX Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Parrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagers K.,
 RA Krogh A., Larsen T.S., Leach S., Moulé S., O'Gaora P., Parry C.,
 RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrrell B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA DOI=10.1128/JB.185.7.2330-2337.2003;
 RG Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 RN [3]
 RP NUCLEOTIDE SEQUENCE OF 1-74.
 RC STRAIN=IMSS-1;
 RX Fernandez-Mora M., Calva E.;
 RA Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
 L-glutamate.
 CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the class-I pyridoxal-phosphate-dependent
 aminotransferase family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.
 CC -----
 DR EMBL; AL627268; CAD053398.1; -; Genomic DNA.
 DR EMBL; AE016840; AAO69551.1; -; Genomic DNA.
 DR EMBL; X89757; CAA61906.1; -; Genomic DNA.
 DR HSSP; P00509; IART.
 DR SMR; O56114; 1-396.
 DR InterPro; IPR004839; Aminotrans_I/II.
 DR InterPro; IPR000796; Asp_trans.
 DR InterPro; IPR004838; Nitransf_1_BS.
 DR PANTHER; PTHR11879; Asp_trans; 1.
 DR Pfam; PF00155; Aminotran_1_2; 1.
 DR PRINTS; PR00799; TRANSAMINASE.
 DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
 DR AMINOTRANSFERASE; Complete proteome; Pyridoxal phosphate; Transferase.
 FT ACT SITE 374
 FT BINDING 246 246 Pyridoxal phosphate (covalent) (By
 similarity).
 FT SEQUENCE 396 AA; 43507 MW; 974C1585438D02FB CRC64;
 Query Match 96.7%; Score 1978; DB 1; Length 396;
 Best Local Similarity 95.7%; Pred. No. 3.5e-141;
 Matches 379; Conservative 12; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MFENITAPADPILGLADLFRADRPCKINIGVYKDETKPTVLTSVKKAEQVLLNE 60
 DB 1 MFENITAPADPILGLADLFRADRPCKINIGVYKDETKPTVLTSVKKAEQVLLNE 60

QY 61 TTKNYLGIDGIPERGRCTOELLFGKGSALINDKXARTAOBTGGTCALRVAADFLAKNTSV 120
 DB 61 TTKNYLGIDGIPERGRCTOELLFGKGSALINDKXARTAOBTGGTCALRVAADFLAKNTSV 120
 QY 121 KRVVSNPSPNPKHSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
 DB 121 KRVVSNPSPNPKHSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
 QY 181 CHNPTGIDPTLEQMQVLAELSVEKGLPLPFDFAVYOGFARGLEEDAEGLRAFAALHKLIV 240
 DB 181 CHNPTGIDPTLEQMQVLAELSVEKGLPLPFDFAVYOGFARGLEEDAEGLRAFAALHKLIV 240
 QY 241 ASSYKNGFLYNERVAGCTLVADSETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
 DB 241 ASSYKNGFLYNERVAGCTLVADSETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
 QY 301 DALRAIWEQELTDMRQIQRMQRQLFVNTLQSKGNRDFSFIIKONGMFSFGLTKEQVLR 360
 DB 301 DALRAIWEQELTDMRQIQRMQRQLFVNTLQSKGNRDFSFIIKONGMFSFGLTKEQVLR 360
 QY 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
 DB 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 8

QSPG88 SALPA
 ID QSPG88 SALPA PRELIMINARY; PRT; 396 AA.
 AC QSPG88;
 DT 01-FEB-2005 (TREMBLrel. 29, Created)
 DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
 DE Aspartate aminotransferase.
 GN Name=aspC; OrderedLocusNames=SPA1800;
 OS Salmonella paratyphi-a
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=54388;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 9150;
 RX PubMed=15531882; DOI=10.1038/ng1470;
 RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
 RA Porwollik S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,
 RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
 RA Kohlberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
 RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
 RA Delehaanty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
 RA Spieth J., Wilson R.K.;
 RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
 restricted serovars of Salmonella enterica that cause typhoid.";
 RL Nat. Genet. 36:1268-1274(2004).
 DR EMBL; CP000026; AAV77716.1; -; Genomic_DNA.
 DR SMR; QSPG88; 1-396.
 DR GO; GO:0008483; F:transaminase activity; IEA.
 DR GO; GO:0006520; P:amino acid metabolism; IEA.
 DR GO; GO:0009058; P:biosynthesis; IEA.
 DR InterPro; IPR004839; Aminotrans_I/II.
 DR InterPro; IPR000796; Asp_trans.
 DR InterPro; IPR004838; Nitransf_1_BS.
 DR Pfam; PF00155; Aminotran_1_2; 1.
 DR PRINTS; PR00799; TRANSAMINASE.
 DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
 DR AMINOTRANSFERASE; Complete proteome; Transferase.
 KW AMINOTRANSFERASE; Complete proteome; Transferase.
 SQ SEQUENCE 396 AA; 43549 MW; 88E9AEDEEA9D149C CRC64;
 Query Match 96.7%; Score 1978; DB 2; Length 396;
 Best Local Similarity 95.7%; Pred. No. 3.5e-141;
 Matches 379; Conservative 11; Mismatches 6; Indels 0; Gaps 0;
 QY 1 MFENITAPADPILGLADLFRADRPCKINIGVYKDETKPTVLTSVKKAEQVLLNE 60
 DB 1 MFENITAPADPILGLADLFRADRPCKINIGVYKDETKPTVLTSVKKAEQVLLNE 60

Db 1 MFENITAAPADPILGLADLFRADDRPGKINLGIVGYKDETGTPTVLTSVKKAHQYLENE 60
 QY 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTATQGTGALRVAADFLAKNTSV 120
 Db 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTATQGTGALRVAADFLAKNTSV 120
 QY 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDPDALINSNEAQAGDVLVFGHC 180
 Db 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDPDALINSNEAQAGDVLVFGHC 180
 QY 181 CHNPTGIDPTLEOWTFLAQLSVKQWLPFLDFAYOGFARGLEDAEGLRFAAMHKLIV 240
 Db 181 CHNPTGIDPTLEOWTFLAQLSVKQWLPFLDFAYOGFARGLEDAEGLRFAAMHKLIV 240
 QY 241 ASSYKSNFGLYNERVGACTLVAADSETVDRAFQSMKAAIRANYSNPPAHGASVATILSN 300
 Db 241 ASSYKSNFGLYNERVGACTLVAADSETVDRAFQSMKAAIRANYSNPPAHGASVATILSN 300
 QY 301 DALRAIWEQELTDMQRIOQMRLQVNTLQKGANRDFSFIKQGMFSGLTKEQVLR 360
 Db 301 DVLRAIWEQELTDMQRIOQMRLQVNTLQKGANRDFSFIKQGMFSGLTKEQVLR 360
 QY 361 LREBFVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
 Db 361 LREBFVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 9

Q6D451_ERWCT PRELIMINARY; PRT; 396 AA.
 ID Q6D451_ERWCT PRELIMINARY; PRT; 396 AA.
 AC Q6D451;
 DT 25-OCT-2004 (TremBLrel. 28, Created)
 DT 25-OCT-2004 (TremBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TremBLrel. 28, Last annotation update)
 DE Aspartate aminotransferase (EC 2.6.1.1).
 GN Name=aspC; OrderedLocusNames=ECA2543;
 OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Pectobacterium.
 OX NCBI_TaxID=29471;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=SCRI 1043 / ATCC BAA-672;
 RX PubMed=15263089; DOI=10.1073/pnas.0402421011;
 RA Bell K.S., Sebailia M., Pritchard L., Holden M.T.G., Hyman L.J.,
 RA Holeva M.C., Thomson N.R., Brooks S.D., Churcher L.J.C., Mungall K.,
 RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
 RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
 RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
 RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
 RT "Genome sequence of the enterobacterial phytopathogen Erwinia
 RT carotovora subsp. atroseptica and characterization of virulence
 RT factors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
 DR EMBL; BX950851; CAG75442.1; -, Genomic_DNA.
 DR SMR; Q6D451; 1-396.
 DR GO; GO:0004069; F:aspartate transaminase activity; IEA.
 DR GO; GO:0006520; P:amino acid metabolism; IEA.
 DR GO; GO:0009058; P:biosynthesis; IEA.
 DR InterPro; IPR004839; Aminotrans_I/I.
 DR InterPro; IPR000796; Asp trans.
 DR Pfam; PF00155; Aminotran_1_2; 1.
 DR PRINTS; PR00799; TRANSAMINASE.
 DR PROSITE; PS00105; AA TRANSFER_CLASS_1; 1.
 KW Aminotransferase; Complete proteome; Transferase.
 SQ SEQUENCE 396 AA; 43081 MW; 817240F2834105A7 CRC64;

Query Match 87.0%; Score 1780; DB 2; Length 396;
 Best Local Similarity .84.6%; Pred. No. 3.4e-126;
 Matches 335; Conservative 29; Mismatches 32; Indels 0; Gaps 0;

QY 1 MFENITAAPADPILGLADLFRADDRPGKINLGIVGYKDETGTPTVLTSVKKAHQYLENE 60

Db 1 MFENITAAPADPILGLADLFRADDRPGKINLGIVGYKDETGTPTVLTSVKKAHQYLENE 60
 QY 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTATQGTGALRVAADFLAKNTSV 120
 Db 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTATQGTGALRVAADFLAKNTSV 120
 QY 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDPDALINSNEAQAGDVLVFGHC 180
 Db 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDPDALINSNEAQAGDVLVFGHC 180
 QY 181 CHNPTGIDPTLEOWTFLAQLSVKQWLPFLDFAYOGFARGLEDAEGLRFAAMHKLIV 240
 Db 181 CHNPTGIDPTLEOWTFLAQLSVKQWLPFLDFAYOGFARGLEDAEGLRFAAMHKLIV 240
 QY 241 ASSYKSNFGLYNERVGACTLVAADSETVDRAFQSMKAAIRANYSNPPAHGASVATILSN 300
 Db 241 ASSYKSNFGLYNERVGACTLVAADSETVDRAFQSMKAAIRANYSNPPAHGASVATILSN 300
 QY 301 DALRAIWEQELTDMQRIOQMRLQVNTLQKGANRDFSFIKQGMFSGLTKEQVLR 360
 Db 301 DALRAIWEQELTDMQRIOQMRLQVNTLQKGANRDFSFIKQGMFSGLTKEQVLR 360
 QY 361 LREBFVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
 Db 361 LREBFVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 10

Q8ZG95_YERPE PRELIMINARY; PRT; 396 AA.
 ID Q8ZG95_YERPE PRELIMINARY; PRT; 396 AA.
 AC Q8ZG95;
 DT 01-MAR-2002 (TremBLrel. 20, Created)
 DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
 DE Aspartate aminotransferase (EC 2.6.1.1).
 GN Name=aspC; OrderedLocusNames=YPO1410;
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
 RA Parkhill J., Wren B.W., Thomson N.R., Titchell R.W., Holden M.T.G.,
 RA Prentice M.B., Sebailia M., James K.D., Churcher C.M., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 DR EMBL; AJ414148; CAC90239.1; -, Genomic_DNA.
 DR PIR; AD0172; AD0172.
 DR HSP; Q8ZG95; 1-396.
 DR SMR; Q8ZG95; 1-396.
 DR GO; GO:0004069; F:aspartate transaminase activity; IEA.
 DR GO; GO:0006520; P:amino acid metabolism; IEA.
 DR GO; GO:0009058; P:biosynthesis; IEA.
 DR InterPro; IPR004839; Aminotrans_I/I.
 DR InterPro; IPR000796; Asp trans.
 DR Pfam; PF00155; Aminotran_1_2; 1.
 DR PRINTS; PR00799; TRANSAMINASE.
 DR PROSITE; PS00105; AA TRANSFER_CLASS_1; 1.
 KW Aminotransferase; Complete proteome; Transferase.
 SQ SEQUENCE 396 AA; 43118 MW; F679902DC5EE9256 CRC64;

Query Match 86.8%; Score 1775; DB 2; Length 396;
 Best Local Similarity 86.1%; Pred. NO. 8.1e-126;
 Matches 341; Conservative 16; Mismatches 39; Indels 0; Gaps 0;

```

QY 1 MFENITAAPADPILGLADLFRADERPGKINLGIVGYKDETKTPTVLTSVKKAEQVYLLENE 60
DB 1 MFEKITAAPADPILGLTDIFRADDDRAHKINLGIVGYKDETKTPTVLTSVKKAEQVYLLENE 60
QY 61 TTKNYLIGDGIPEGRCTQELLFGKGSALINDKARTAOPTGGTGALRVAADFLAKNTSV 120
DB 61 ATKNYLIGDGLVPFASCTQELLFGANSIIADKARTAOPTGGTGGLRIAADFTAHQTS 120
QY 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
DB 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
QY 181 CHNPTGIDPTLEQWOTLAQLSVKGLPLDFAYQGFARGLEEDAEGLRAFAAMHKLIV 240
DB 181 CHNPTGIDPTLEQWOTLAQLSVKGLPLDFAYQGFARGLEEDAEGLRIFAATHQELIV 240
QY 241 ASSYSKFNGLYNERVGACTLVAADSETVDRAFSPQKAAIRANYSNPPAHGASVVATILSN 300
DB 241 CSSYSKFNGLYNERVGACTLVAADSNVADTAFSQKAVIRANYSNPPAHGASVVATILSN 300
QY 301 DALRAIWEQELTDMRQIRQMRQLFVNTLQKGNRDFSFIIKQNGMFSFSLTKKEQVLR 360
DB 301 AALRAIWEQELTDMRQIRQMRQLFVNTLQKGAQDFSFIIQNGMFSFSLTKKEQVLR 360
QY 361 LREEFGVYAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
DB 361 LRDEFVAVVNSGRVNVAGMTPDNNAPLCEAIVAVL 396

```

RESULT 11

```

ID Q66CG9 YERPS PRELIMINARY; PRT; 396 AA.
AC Q66CG9
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Aspartate aminotransferase (EC 2.6.1.1).
GN Name=aspC; OrderedLocusNames=YPTB1434;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Verges L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis."
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX936398; CAH20674.1; -; Genomic_DNA.
DR SMR; Q66CG9: 1-396.
DR GO; GO:0004069; F:aspartate transaminase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR004839; Aminotrans_I/II.
DR InterPro; IPR000796; Asp_trans.
DR InterPro; IPR004838; Ntrtrans_1.
DR Pfam; PF00155; Aminotran_1_2; 1.
DR PRINTS; PR00799; TRANSAMINASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
KW Aminotransferase; Complete proteome; Transferrase.
SQ SEQUENCE 396 AA; 43118 MW; F679902DC5EE9256 CRC64;

```

Query Match 86.8%; Score 1775; DB 2; Length 396;
 Best Local Similarity 86.1%; Pred. No. 8.1e-126;
 Matches 341; Conservative 16; Mismatches 39; Indels 0; Gaps 0;

```

QY 1 MFENITAAPADPILGLADLFRADERPGKINLGIVGYKDETKTPTVLTSVKKAEQVYLLENE 60
DB 1 MFEKITAAPADPILGLTDIFRADDDRAHKINLGIVGYKDETKTPTVLTSVKKAEQVYLLENE 60
QY 61 TTKNYLIGDGIPEGRCTQELLFGKGSALINDKARTAOPTGGTGALRVAADFLAKNTSV 120
DB 61 ATKNYLIGDGLVPFASCTQELLFGANSIIADKARTAOPTGGTGGLRIAADFTAHQTS 120
QY 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
DB 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
QY 181 CHNPTGIDPTLEQWOTLAQLSVKGLPLDFAYQGFARGLEEDAEGLRAFAAMHKLIV 240
DB 181 CHNPTGIDPTLEQWOTLAQLSVKGLPLDFAYQGFARGLEEDAEGLRIFAATHQELIV 240
QY 241 ASSYSKFNGLYNERVGACTLVAADSETVDRAFSPQKAAIRANYSNPPAHGASVVATILSN 300
DB 241 CSSYSKFNGLYNERVGACTLVAADSNVADTAFSQKAVIRANYSNPPAHGASVVATILSN 300
QY 301 DALRAIWEQELTDMRQIRQMRQLFVNTLQKGNRDFSFIIKQNGMFSFSLTKKEQVLR 360
DB 301 AALRAIWEQELTDMRQIRQMRQLFVNTLQKGAQDFSFIIQNGMFSFSLTKKEQVLR 360
QY 361 LREEFGVYAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
DB 361 LRDEFVAVVNSGRVNVAGMTPDNNAPLCEAIVAVL 396

```

RESULT 12

```

ID Q8D050 YERPE PRELIMINARY; PRT; 401 AA.
AC Q8D050; 074VU3;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE Aspartate aminotransferase.
GN Name=aspC; OrderedLocusNames=YPI1183, Y2760;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=91001;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
RT avirulent to humans."
RL DNA Res. 11:179-197(2004).
DR EMBL; AE013879; AAM86312.1; -; Genomic_DNA.
DR EMBL; AE017131; AAS61426.1; -; Genomic_DNA.
DR HSSP; P00509; IART.
DR SMR; Q8D050; 6-401.
DR GO; GO:0008483; F:transaminase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR004839; Aminotrans_I/II.

```



```
DR InterPro; IPR000796; Asp_trans.
DR InterPro; IPR004838; NHTtransf_1_BS.
DR Pfam; PF00155; Aminotran_1_2; 1.
DR PRINTS; PR00799; TRANSAMINASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
KW Amino transferase; Transferase.
SQ SEQUENCE 401 AA; 43652 MW; A58AEBB191041831 CRC64;

Query Match      86.1%; Score 1775; DB 2; Length 401;
Best Local Similarity 86.1%; Pred. No. 8.3e-126;
Matches 341; Conservative 16; Mismatches 39; Indels 0; Gaps 0;

QY 1 MFENITAAPADPILGLADLFRADRPKGKINIGVYKDETKTPTVLTSVKKAQVLLNE 60
DB 6 MFEKITAAPADPILGLTDFRADRAHKNILGIGVYKDETKTPTVLTSVKKAQVLLNE 65
QY 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTAQTPGGTGALRVAADFIAKNTSV 120
DB 66 ATKNYLGIDGLPVFASCTQELLFGANSALIIADKARTAQTPGGTGGLRIAADFIAHQ TSA 125
QY 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
DB 126 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 185
QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGLPLFPAYQGFARGLEEDAEGLRAFAAMHKLIV 240
DB 186 CHNPTGIDPTLEQWOTLAQLSVEKGLPLFPAYQGFARGLEEDAEGLRAFAAMHKLIV 245
QY 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
DB 246 CSSYSKNFGLYNERVGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVATILSN 305
QY 301 DALRAIWEQELTDMRQRIQRMQLFVNTLQKGNRDFSFIIKONGMFSFGLTKQVLR 360
DB 306 AALRAIWEQELTDMRQRIQRMQLFVNTLQKGNRDFSFIIKONGMFSFGLTKQVLR 365
QY 361 LREEFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 396
DB 366 LRDEFVAVVNSGRVNVAGMTDPNNAPLCEAIVAVL 401

RESULT 13
Q7N625 PHOLL PRELIMINARY; PRT; 396 AA.
AC Q7N625
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Aspartate aminotransferase (Transaminase A).
GN Name=aspC; OrderedLocusNames=plu1750;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
RA Duchaud E., Rusniok K., Frangeul L., Buchrieser C., Givaudan A.,
RA Taourit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dagaia E., Detose R., Deruelle S., Freyssonnet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Siquier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens."
RL Nat. Biotechnol. 21:1307-1313(2003).
DR EMBL; BX571864; CAE14043.1; -; Genomic_DNA.
DR HSSP; P00509; 1AET.
DR SMR; Q7N625; 1-396.
DR PhotocList; plu1750; -.
DR GO; GO:0008483; F:transaminase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
```

```
DR InterPro; IPR004839; Aminotrans_I/II.
DR InterPro; IPR000796; Asp_trans.
DR InterPro; IPR004838; NHTtransf_1_BS.
DR Pfam; PF00155; Aminotran_1_2; 1.
DR PRINTS; PR00799; TRANSAMINASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
KW Amino transferase; Complete proteome; Transferase.
SQ SEQUENCE 396 AA; 43575 MW; 7EC3B64A6DBC626 CRC64;

Query Match      82.2%; Score 1682; DB 2; Length 396;
Best Local Similarity 79.8%; Pred. No. 8.9e-119;
Matches 316; Conservative 38; Mismatches 42; Indels 0; Gaps 0;

QY 1 MFENITAAPADPILGLADLFRADRPKGKINIGVYKDETKTPTVLTSVKKAQVLLNE 60
DB 1 MFEKITAAPADPILGLADSFSDPRTNKNILGIGVYKDETKTPTVLTSVKKAQVLLNE 60
QY 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTAQTPGGTGALRVAADFIAKNTSV 120
DB 61 TTKNYLGIDGLPVFASCTQELLFGKDHVPVVDKARTAQSPGGTGALRIAADFIAKQNA 120
QY 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
DB 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGLPLFPAYQGFARGLEEDAEGLRAFAAMHKLIV 240
DB 181 CHNPTGIDPTLEQWOTLAQLSVEKGLPLFPAYQGFARGLEEDAEGLRIFAKNHNLIV 240
QY 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
DB 241 ASSYSKNFGLYNERVGACTIVASDSTAFAKAFSAQAKAIRANYSNPPAHGASIVTILSN 300
QY 301 DALRAIWEQELTDMRQRIQRMQLFVNTLQKGNRDFSFIIKONGMFSFGLTKQVLR 360
DB 301 EDLKAWEQELTDMRQRIQRMQLFVNTLQKGNRDFSFIIKONGMFSFGLTKQVLR 360
QY 361 LREEFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 396
DB 361 LRDEFVAVVNSGRVNVAGMTDPNNAPLCEAIVAVL 396

RESULT 14
Q5DI91_9PAST PRELIMINARY; PRT; 396 AA.
AC Q5DI91_9PAST
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE AspC.
GN Name=aspC;
OS Actinobacillus porcitonisillarum.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=189834;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CCUG 46996;
RX PubMed=15863281; DOI=10.1016/j.vetmic.2005.01.020;
RA Kuhnert P., Schlatter Y., Frey J.;
RT "Characterization of the type I secretion system of the RTX toxin
RT AptII in Actinobacillus porcitonisillarum."
RL Vet. Microbiol. 107:225-232(2005).
DR EMBL; AY956600; AA21413.1; -; Genomic_DNA.
DR InterPro; IPR004839; Aminotrans_I/II.
DR InterPro; IPR000796; Asp_trans.
DR Pfam; PF00155; Aminotran_1_2; 1.
DR PRINTS; PR00799; TRANSAMINASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
SQ SEQUENCE 396 AA; 43374 MW; 605EA13BFCDF4BD0 CRC64;

Query Match      66.0%; Score 1349; DB 2; Length 396;
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 18:00:15 ; Search time 47 Seconds
(without alignments)
696.587 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045

Sequence: 1 MFENITAAPADPILGLADLF.....VAGMTPDNMAPLCEAIVAVL 396

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/6_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/H_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/PGTUS_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/RE_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2045	100.0	396	2	US-08-985-908-24
2	1931	94.4	402	2	US-09-489-039A-12073
3	982	48.0	413	2	US-09-252-991A-28999
4	909.5	44.5	543	2	US-09-252-991A-29078
5	908	44.4	233	2	US-09-543-681A-7977
6	871.5	42.6	402	2	US-09-540-236-2235
7	844.5	41.3	445	2	US-09-489-039A-8472
8	843.5	41.2	419	2	US-09-543-681A-7295
9	830.5	40.6	427	2	US-09-328-352-4895
10	794	38.8	430	2	US-09-976-594-430
11	756	37.0	416	2	US-09-949-016-7087
12	726.5	35.5	395	2	US-09-198-452A-790
13	726.5	35.5	395	2	US-09-438-185A-742
14	694	33.9	365	2	US-09-801-874-5
15	542	26.5	303	2	US-09-248-796A-17477
16	486.5	23.8	289	2	US-09-248-796A-17478
17	436.5	21.3	141	2	US-09-543-681A-7889
18	359.5	17.6	380	2	US-09-801-874-2
19	321.5	15.7	241	2	US-09-248-796A-17480
20	284.5	13.9	188	2	US-09-248-796A-17479
21	233.5	11.4	254	2	US-09-801-874-4
22	191.5	9.4	123	2	US-09-513-999C-6182
23	182	8.9	115	2	US-09-248-796A-17476
24	161.5	7.9	434	2	US-09-134-001C-4695
25	153.5	7.5	429	2	US-09-710-279-3174
26	148	7.2	29	1	US-07-952-817-20
27	148	7.2	29	6	5210025-13

28	121.5	5.9	373	2	US-09-902-540-10929	Sequence 10299, A
29	121.5	5.9	771	2	US-09-712-363-157	Sequence 157, App
30	118.5	5.8	413	2	US-09-489-039A-13270	Sequence 13270, A
31	117	5.7	401	2	US-09-252-991A-27198	Sequence 27198, A
32	110.5	5.4	345	2	US-09-724-623-65	Sequence 65, Appl
33	109.5	5.4	378	2	US-09-489-039A-10110	Sequence 10110, A
34	108.5	5.3	393	2	US-09-543-681A-4673	Sequence 4673, Ap
35	108	5.3	843	1	US-09-172-977-3	Sequence 3, Appli
36	108	5.3	843	2	US-09-404-108-3	Sequence 3, Appli
37	107	5.2	1839	1	US-09-172-977-4	Sequence 4, Appli
38	107	5.2	1839	2	US-09-404-108-4	Sequence 4, Appli
39	107	5.2	2753	2	US-09-949-016-7659	Sequence 7659, Ap
40	107	5.2	2753	2	US-09-949-016-7660	Sequence 7660, Ap
41	107	5.2	3924	2	US-09-538-092-1246	Sequence 1246, Ap
42	106.5	5.2	422	2	US-09-252-991A-21571	Sequence 21571, A
43	102	5.0	395	2	US-09-489-039A-14008	Sequence 14008, A
44	102	5.0	551	2	US-09-538-092-738	Sequence 738, App
45	101.5	5.0	889	2	US-09-489-039A-7241	Sequence 7241, Ap

ALIGNMENTS

RESULT 1

US-08-985-908-24

; Sequence 24, Application US/08985908

; Patent No. 6004773

; GENERAL INFORMATION:

; APPLICANT: MASAYUKI ARAKI, MASAKAZU SUGIMOTO, YASUHIKO YOSHIHARA, AND TSUYOSHI N.

; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.

; STREET: 1755 S. JERPERSON DAVIS HIGHWAY, FOURTH FLOOR

; CITY: ARLINGTON

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION NUMBER:

; APPLICATION NUMBER: US/08/985,908

; FILING DATE: 05-DEC-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 8-325659

; FILING DATE: 05-DEC-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: NORMAN F. OBLON

; REGISTRATION NUMBER: 24,618

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-413-3000

; TELEFAX: 703-413-2220

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 396 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-985-908-24

Query Match 100.0%; Score 2045; DB 2; Length 396;

Best Local Similarity 100.0%; Pred. No. 4.2e-205;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADPKGINLIGVYKDETKPVLTSVKAEQYLLENE 60

Db 1 MFENITAAPADPILGLADLFRADPKGINLIGVYKDETKPVLTSVKAEQYLLENE 60

Qy 61 TTKNYLGIDGIPBFGRCCTQELLFCGKSALINDKRAQTAPGGTGAALRVAADFLAKNTSV 120

Fr1 Mar 17 08:04:47 2006

```

; SEQ ID NO 8472
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Kl
US-09-489-039A-8

```

Query Match	41.3%	Score 844.5;	DB 2;	Length 445;
Best Local Similarity	42.1%	Pred. No. 2.2e-79;		
Matches 167; Conservative	74;	Mismatches 155;	Indels 1;	Gaps 1;

Qy	1	MENITAA	PADPIL	GLADL	FRADER	PGKIN	LIGVYK	DET	GKTPVL	TSVKKE	QOYL	LBN	59
Db	49	VFOQ	DAYAGD	PILSMER	FED	PRSD	KVNLSI	GLIYND	GII	IPQ	QVAE	AEARLNAEP	108
Qy	60	ETTKNY	LIGD	IGPE	GRCT	QCELL	FGKS	ALIND	KRART	QTG	GTG	ALRV	119
Db	109	HGAS	LPLW	EGUS	GRQAI	APLL	FGA	HTALK	QNR	IASIQ	TVG	SNAL	168
Qy	120	VKR	VWY	SNP	SWN	HKSV	FNS	AGLE	VRE	YAY	YDA	ENHTL	179
Db	169	ESHV	WVSD	PTW	ENH	IAI	PEG	AGF	EVST	YPM	FDK	ATNG	228
Qy	180	CCNP	TGID	PTL	QEQ	HTLA	QLS	VEK	GWLP	FD	FAY	QCF	239
Db	229	CCNP	TGAD	LTPA	QDR	VV	ELKAR	QLP	IFP	LDIA	OYQ	FGG	288
Qy	240	VASS	YKFN	GLN	ERV	AG	CTIV	AD	SE	TV	DRA	FQMKAA	299
Db	289	VNSF	SKI	PSL	GER	VGG	LV	CC	ED	SETA	GRV	LQ	348
Qy	300	NDA	LRAI	WPE	QELT	DM	RQI	QRM	QRF	VNT	LQ	EKAN	359
Db	349	DAGL	KATW	QAE	VDAM	RAH	LIT	W	RQAL	VDAL	QO	VA	408
Qy	360	RLR	EEFG	YVA	YAS	GRV	NV	AG	MT	PD	NMA	PL	396
Db	409	RLR	DEFG	VTV	LIA	SG	RM	CV	AG	LSR	NV	QO	445

RESULT 8
 US-09-543-681A-7295
 ; Sequence 7295, Application US/09543681A
 ; Patent No. 6605709
 ; GENERAL INFORMATION:
 ; APPLICANT: GARY BRETON
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES AND THERAPEUTICS
 ; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/543.681A
 ; CURRENT FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/128,706
 ; PRIOR FILING DATE: 1999-04-09
 ; NUMBER OF SEQ ID NOS: 8344
 ; SEQ ID NO 7295
 ; LENGTH: 419
 ; TYPE: PRT
 ; ORGANISM: *Proteus mirabilis*
 ; US-09-543-681A-7295

	Query Match	41.2%	Score	843.5;	DB 2;	Length	419;			
	Best Local Similarity	41.9%;	Pred. No.	2.5e-79;						
	Matches	166;	Conservative	75;	Mismatches	154;	Indels	1;	Gaps	1
Qy	1	MFENITAAPADILGLADLFRADERPGKINLGIVYKDERTGKTPTVLTSVKKAEOYL-L	EN	59						
Dd	23	VFQQVEAFAGDPILSLMDVYNKDPQDKINLSIGLYYDEEGKTPILGTVSVARQQLNAMT		82						
Qy	60	ETTKNYLGDIGIPFGRCRTOELLFGKGSAIINDKARTAQTPGGTGALREVAADF	LAKNTS	119						
Dd	83	PTATLYLPMEGLAPYRHVEQTLLFGADNPDIADKKIATIQTUGGSGALKVGDAFLHRVP		142						
Qy	120	VKRVMVSNPSNPNHKSFNVSAGLEVREYAYDAENHTLDFDALINSLEQAQGVVLFHG		179						

Qy		1	MPENTAAPADPILGLADLFRADERPGKINLGIVYKOETGKTPTLVTSVKKAEOYL-LEN	59
D8		23	VFQQVEAFAGDPILSLMDVYNPKDKNLSIGLYDEEGKTPILGTVSARQQLNMT	82
Qy		60	ETTNNYLGDGTFEGRCTOELLFGKGSALINDKRAATAQTGGTGALRVAADFLEAKNTS	119
D8		83	PTATYLPMEGLAPRYHEVOTLLFGADNPDIADKKIATITQTGGSGALKVGADFLHRYFP	142
Qy		120	VKRVMVSNPSFNHHKSFNVSAGLEVREYAYDAENHTTLDFDALINSLNQAQGDDVVLFHG	179

143	SSEWISDPTWDNDHNASIFAGSGFKVNYYPFDPETKGVKFDALLDCFKKPEKSIVLMHP	203
180	CCHNPTGIDPTLEQWQTLAOLSVEKGWMLPUDFAIYQGFARGLEDEAGLRFAFAAMHKELI	239
203	CCHNPTGSDLTAKAQWQDVBILKARQAIPELFIAYQGFABSLDDDDAVARAMAKAGLEPVL	262
240	VASSYKNGFLYNERVCACCTLVAADSETVDRAFQSMKAAIRKANTSNPPAHGASVVATILS	299
263	VSNSFKIFIGYGRAGGLSIVCDNAKECEHVLGQLAGARRIYSSPANYGAQLVNVQLS	322
300	NDALRALWEQELTDMRQIRORMROLFVNTYLQEKGANRDFSFIIKONGHMFSGLTKEQVL	359
323	DHVLTAQWQKEVAHMRDRIKEMRVTLVNAKLEAPEKNFDHLLTORGWFSYTGFSPEQVD	382
360	RLREEFGVYAVASGRVNVAGMTTPDDNMAPLCEAIVAV	395
383	RLREEFGIYLITGRVCMGVNNNNVQRIQAQFAAV	418

RESULT 9

```

US-09-328-352-4895
; Sequence 4895, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND
; FILE OF INVENTION: BAUMANNII FOR DIA
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09328
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4895
; LENGTH: 427
; TYPE: PR1
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4895

```

Query Match 40.6%; Score 830.5; DB 2; Length 427;
 Best Local Similarity 42.1%; Pred. No. 6e-78;
 Matches 167; Conservative 72; Mismatches 157; Indels 1; Gaps 1;

[illegible]

RESULT 10
US-09-976-594-430
: Sequence 430, Application US/09976594

```
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 430
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 1273292CD1
US-09-976-594-430

Query Match          38.8%; Score 794; DB 2; Length 430;
Best Local Similarity 41.2%; Pred. No. 4e-74;
Matches 161; Conservative 67; Mismatches 161; Indels 2; Gaps 2;

Qy      4 NITAAPADPILGLADLRADERPKINLGIGVYKDETKTPVLTSVKKAEOYLLNETTK 63
       :|::||:::||:|::||:|::||:|::||:|::||:|::||:|::||:|::||:|
Db      35 HVEMGPPDPILGVTEAFKRDTNSKKMNLGVAYRDDNGKPYVLPSPVRKAEQAIAKNLDK 94

Qy      64 NYLGIDIGPEGRCTOELLFKGSGALINDKEARTAOTPGGTGALRVAADFIAKNTSVKR- 122
       :|::||:::||:|::||:|::||:|::||:|::||:|::||:|::||:|
Db      95 EYLPIGSLAEFCASALALGENSEVLKSGRFVTQTISGTGALRIGASFLOREFKFSRD 154

Qy      123 VWTSNRPWNHKSFVENSAGLEVRVAYYAENHTLDFDALINSNEAQAGDVVLFHGCGH 182
       :|::||:::||:|::||:|::||:|::||:|::||:|::||:|::||:|
Db      155 VFLEPKFTWGNHTPLFRDGMQLQGYRYDPTKCFDFTGAVEDISKIPEQSIVLLHCAH 214

Qy      183 NPTEGIDPLEQMOTLAQLSVEKGWLPFLDFAYQGFARG-LEEDAEGLRAFAAMEHELIVA 241
       :|::||:::||:|::||:|::||:|::||:|::||:|::||:|::||:|
Db      215 NPTGVDPRPEQWEKIATVVKKRNLFAPFDMAYQGFGSDGDKDAWAVRHFTIEQGINVCLC 274

Qy      242 SSVSKNFGLYNERVGACTLVAADSETVDRAFSSQKAAIRANYSNPPAHGASVATILSND 301
       :|::||:::||:|::||:|::||:|::||:|::||:|::||:|::||:|
Db      275 QSAKNNGLYGERVGAFTMWCKDAEKRVESOLKLIRPMYSNPPLNGARIAAAAILNTP 334

Qy      302 ALRAIWQEQLTDMQRQWRQLFWNTIQEKGANRDFSFIKONGMFSGLTKEQVRL 361
       :|::||:::||:|::||:|::||:|::||:|::||:|::||:|::||:|
Db      335 DLRKQWLQEVKGMADRIGMTQVLVSNLKKEGSTHNWOHITDQIGMFCFTGLKPEQVERL 394

Qy      362 REEFQYAVASGRVNAGWTPDDNMAPLCEAI 392
       :|::||:::||:|::||:|::||:|::||:|::||:|::||:|::||:|
Db      395 IKESFYMTXDRGISVAGVTSNVGYLAHAH 425

RESULT 11
US-09-949-016-7087
; Sequence 7087, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SEQ ID NO 7087
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7087

Query Match      37.0%; Score 756; DB 2; Length 416;
Best Local Similarity 43.0%; Pred. No. 3.6e-70;
Matches 166; Conservative 54; Mismatches 156; Indels 10; Gaps 5;

Qy 10 ADPIL--GLADLPRADERPCKINLGIIVGVKDETKGTPVLTSVKKASQYLL--ENETTKNYL 66
Db 16 AQPVLVFKLTADPREDPDPKKNVILGVGARTDDCHPWVLPVVKVKQKANDNSLHMEYL 75

Qy 67 GIDGIEFGRCIOELLFSGKSALINDKARTATQPGTGALRVAADFLAK-----NTSVK 121
Db 76 PILGLAEFRSCASRLALGDDSPALKEKRVGVQSLGTGALRIGADFLARWYNGTNNKNT 135

Qy 122 RVWVSNPWNHKSVPNSAGL--EVRYAYYYDAENHTLDFDALINSLEAQAQGVVLFHGC 180
Db 136 PVTVSSPTWENHNNAVFAAGFKDIRSYRYWDAEKRGDLQGGFLNDLENAPESIVVLHAC 195

Qy 181 CHNPTGIDPTLEOWQTLAQLSVEKGMWLPDFDAYQGFARG--LEEDAEGLRAPAAHKELI 239
Db 196 AHNPTGIDPTPEQWKQIALSVMKHFRLFPFSDSAYQGFASGNLERDWAIRYFVSEGFEEF 255

Qy 240 VASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVAATILS 299
Db 256 CAQSFKNFGLYNERVGNLTIVGKPEFSLQIVLSQMEKIVRITWSNPPAQGARIVASTLS 315

Qy 300 NDALRAIWQBELTDMQRQIORMRQLFWNTLQEKANRDPFIIKQNGMPSFSGLTKEQVL 359
Db 316 NPELFEWTGNVXTMADRILTNRSLEARLEALKTPGTWNHITDQICMFSTGLNPKQVE 375

Qy 360 RLREEFQVAVASGRVNVAGMTPDNM 385
Db 376 YLYNEKHIVLLPSGRINVSLGTTKNL 401

RESULT 12
US-09-198-452A-790
; Sequence 790, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides,
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 790
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-790

Query Match      35.5%; Score 726.5; DB 2; Length 395;
Best Local Similarity 39.6%; Pred. No. 4e-67;
Matches 156; Conservative 75; Mismatches 154; Indels 9; Gaps 3;

Qy 2 FENITAAPADPILGLADLFRADERPCKINLGIIVGVKDETKGTPVLTSVKKASQYLLNET 61
Db 4 FNIHPTSPDAILGLQNVPFAKRPCKNVLGVVHEHPQKRVGGISCIKRAQTVILEEQ 63

Qy 62 TKNYLGDIGDIPFGRCIOELLFG--KGSALINDKARTATQPGTGALRVAADFLAKNTS 119
Db 64 NKSYPILSGIQIFLDEMRELVFQAVDPSAIV-----FQSLGGTGALHGLARLLSVAKG 117

Qy 120 VKRWVSNPWNHKSVPNSAGLEVREYAYYDAENHTLDFDALINSLEAQAQGVVLFHGC 179
Db 118 SGKVYVPEQTWSNHIRFISQEGLEVTRYPVYSKEQKQLLFEPLIAPLKEVKNNSVILLHG 177

```

```

Qy 180 CCHNPTGIDPTLEQWQTIAQLSVEKGWLPLFDFAFYQGFAKGLBEDAEGLPAPAMHKELI 239
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
Db 178 CCHNPTGVDFTEDMWKELAILMKERLIPFFDTAYOQGFHAGHIELDRKPFIIFISEGNTVL 237
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::

Qy 240 VASSYSKNFGLNYRVGACTIVAADSEVTVDRAFQMKAAARANYSNPPAHGASVWATILS 299
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
Db 238 VAASSCKNFALCERVGYFAVHSITFTDELVKIHSFLEEKIRGEYSSPQRGWVEIVSTILLS 297
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::

Qy 300 NDALARIMWEQUETDMWRQIQRMRLQFLVNTLOEKGANRDFFSIKQNGMFSFSLGTKEQVL 359
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
Db 298 NPYLKEEMOSELNFIRESLGKMRTFRVALR-KVAGHTFDFLLSQHGFFAYPGFSDKQVL 356
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::

Qy 360 RLREBEGVVAVASGRVNNAVGMTDPMAPLCEAIv 393
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::
Db 357 FLREQHAVTYTAGRMNLNGITEKNIHDHVQSFI 390
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||::

RESULT 13
US-09-438-185A-742
; Sequence 742, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108, 279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128, 606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 742
; LENGTH: 397
; TYPE: PRt
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CpN0740
US-09-438-185A-742
```

Qy	360	RLREBFGYVAVASGRVNVAGTMDPNMAPLCEAIV	392
Db	359	FLREQHAVTTAGGRKMLNGITEKNIDHVQSF	392
RESULT 14			
US-09-801-874-5			
; Sequence 5, Application US/09801874			
; Patent No. 6582935			
; GENERAL INFORMATION:			
; APPLICANT: YAN, Chunhua et al.			
; TITLE OF INVENTION: ISOLATED HUMAN AMINOTRANSFER-			
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLEC-			
; TITLE OF INVENTION: AMINOTRANSFERASE PROTEINS,			
; TITLE OF INVENTION: AND USES THEREOF			
; FILE REFERENCE: CL000615			
; CURRENT APPLICATION NUMBER: US/09/801.874			
; CURRENT FILING DATE: 8001-03-09			
; NUMBER OF SEQ ID NOS: 5			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 5			
; LENGTH: 365			
; TYPE: PRT			
; ORGANISM: Human			
US-09-801-874-5			
Query Match 33.9%; Score 694; DB 2;			
Best Local Similarity 43.2%; Pred. No. 8.9e-64;			
Matches 153; Conservative 47; Mismatches 144			
Qy	10	ADPIL--GLADLFRADRPFGKINLIGIVYKDETCKTPVP	
Db	9	AQPVLVFKLTADFREDPPRKVNGLGVGNRYRTDDCHPWV	
Qy	67	GDGIGPEFGCRCTQELLFGKGSALINDKRAFTAQTPGCTG	
Db	69	PILGLAEFFSCASRLALGDSDPALKEKVGCVGLSGTGTG	
Qy	122	RWVSNPSPWNHKSVPNSAGL-EVREYAYYDAENHTLD	
Db	129	PVVYSSPTWENHNNAVFSAAAGFKDTSRYWDAEKRGGLD	
Qy	181	CNPCTGIDPTLSOWTQTLAQLSVEKGWLPFLDFAYQGFAP	
Db	189	AHNPCTGIDPTPQWQIASVWKRPFLPFFFDSDAQQGFAP	
Qy	240	VASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKKK	
Db	249	CAQSFSKNFGLYNERVGNLTVVGKEPSILQVLVSQMEK	
Qy	300	NDALRAINWQELTDHQRIOIRMQLFVNTLQEKGANRDR	
Db	309	NPELFEFTGNVKTWADRIILTMRSRLRARLEALKTPGGT	
RESULT 15			
US-09-248-796A-17477			
; Sequence 17477, Application US/09248796A			
; Patent No. 6747137			
; GENERAL INFORMATION:			
; APPLICANT: Keith Weinstein et al			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID S			
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEU			
; FILE REFERENCE: 107196.132			
; CURRENT APPLICATION NUMBER: US/09/248.796A			
; CURRENT FILING DATE: 1999-02-12			
; PRIOR APPLICATION NUMBER: US 60/074.725			
; PRIOR FILING DATE: 1998-02-13			
; PRIOR APPLICATION NUMBER: US 60/096.409			
; PRIOR FILING DATE: 1998-08-13			
; NUMBER OF SEQ ID NOS: 28208			
; SEQ ID NO 17477			


```

; LENGTH: 303
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17477

Query Match      26.5%; Score 542; DB 2; Length 303;
Best Local Similarity 43.5%; Pred. No. 5.3e-48;
Matches 114; Conservative 35; Mismatches 101; Indels 12; Gaps 3;

QY      2 FENITAAPADPILGLADLFRADERPGKINLIGIVYKDETGTPTVLTSVKKAQYLLLENET 61
Db      : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : |||
36 WNDIPLAPDKILGISEAYNNDSPQKINLGVGYARDNSGKPIIPPSVKKAEIILLGKET 95
QY      62 TKNYLGIDGIPFGRCTOELLF-----GKGSALINDKRAKTAQTPGGTGALRVAADFIA 115
Db      : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : |||
96 EKEYTAIVGSKNFQSVKRNFIENNSKNDANGQLIDDDRIVTAQTISGTGSLRVIADFLN 155
QY      116 KNTSVKRVVWVSNPSVPHKSVFNSAGLEVREYAYDAENHTLDPDALINSLNEAQAGDVV 175
Db      : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : |||
156 RFSYNNKILVPKPTWANHVAVFKDAGLEPEFYFYYTETSKNDLDYANLKKSLTAAPEGSIV 215
QY      176 LFHGGCHNPTGIDPTLEQWOTLAQLSVEKGWLPDFDFAYQGFARGLE-EDAEGLRFAFAAM 234
Db      : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : |||
216 LLHACHNPTGMDLISEQWDEVLIQVQDKKFPFLVDMAYQGFASGKPFEDIGLIRKLTKL 275
QY      235 HKE-----LIVASSYSKNGFLY 251
Db      : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : |||
276 ANENKIPSFALCQSPAKNGGLY 297

```

Search completed: March 14, 2006, 18:01:38
Job time : 48 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 18:11:25 ; Search time 166 Seconds
(without alignments)
996.749 Million cell updates/sec

Title: US-10-673-786A-2
Perfect score: 2045
Sequence: 1 MFENITRAPADPILGLADLF.....VAGTMDNMAPLCEAIVAVL 396

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing, first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2045	100.0	396	4	US-10-369-493-785
2	2045	100.0	396	4	US-10-673-786A-2
3	1654	80.9	396	4	US-10-369-493-21125
4	1318	64.4	322	4	US-10-369-493-231
5	1311	64.1	397	4	US-10-275-026A-188
6	1005.5	48.2	394	4	US-10-369-493-11824
7	996	48.7	396	4	US-10-369-493-7497
8	996	48.7	398	4	US-10-369-493-4738
9	982	48.0	397	4	US-10-369-493-8424
10	975	47.7	397	4	US-10-369-493-15828
11	975	47.7	397	4	US-10-369-493-16208
12	975	47.7	398	4	US-10-369-493-15460
13	947	46.3	398	4	US-10-369-493-14025
14	924	45.2	395	4	US-10-369-493-9458
15	917	44.8	395	4	US-10-369-493-9222
16	917	44.8	420	4	US-10-425-114-63613
17	914	44.7	420	4	US-10-425-114-63242
18	914	44.7	459	4	US-10-767-701-46300
19	913	44.6	423	4	US-10-425-114-57028
20	911.5	44.6	395	4	US-10-369-493-13723
21	911	44.5	417	4	US-10-425-114-68709
22	909	44.4	400	4	US-10-369-493-17487
23	904.5	44.2	464	4	US-10-425-115-197293
24	893	43.7	396	4	US-10-369-493-8659
25	885	43.3	456	4	US-10-424-599-239081
26	884.5	43.3	396	4	US-10-369-493-21154
27	875.5	42.8	396	4	US-10-369-493-287

28	870	42.5	1991	4	US-10-437-963-187401	Sequence 187401, A
29	866	42.3	412	4	US-10-425-115-287905	Sequence 287905, A
30	866	42.3	574	4	US-10-425-115-363667	Sequence 363667, A
31	865	42.3	520	4	US-10-425-115-363669	Sequence 363669, A
32	863	42.2	463	4	US-10-424-599-172508	Sequence 172508, A
33	861	42.1	459	4	US-10-767-701-46309	Sequence 46309, A
34	851	41.6	396	4	US-10-369-493-4827	Sequence 4827, A
35	851	41.6	396	4	US-10-369-493-7586	Sequence 7586, A
36	846	41.4	413	5	US-10-481-032A-290	Sequence 290, A
37	843.5	41.2	397	4	US-10-369-493-23657	Sequence 23657, A
38	831.5	40.7	397	4	US-10-422-366-32	Sequence 32, A
39	831.5	40.7	397	5	US-10-925-216-32	Sequence 32, A
40	831.5	40.7	397	5	US-10-903-582-32	Sequence 32, A
41	831.5	40.7	397	5	US-10-918-401A-32	Sequence 32, A
42	831.5	40.7	397	5	US-10-969-245-32	Sequence 32, A
43	828	40.5	424	6	US-11-097-143-5940	Sequence 5940, A
44	825	40.3	403	5	US-10-739-930-6191	Sequence 6191, A
45	823	40.2	710	5	US-10-450-763-54363	Sequence 54363, A

ALIGNMENTS

RESULT 1
US-10-369-493-785
; Sequence 785, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 785
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-369-493-785

Query Match	100.0%	Score	2045	DB	4	Length	396
Best Local Similarity	100.0%	Pred. No.	4.2e-190				
Matches	396	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MFENITRAPADPILGLADLF	FRADRP	PGKIN	IGVYKDET	GTGKTP	VLTSVKKAEQVLLNE 60
Db	1	MFENITRAPADPILGLADLF	FRADRP	PGKIN	IGVYKDET	GTGKTP	VLTSVKKAEQVLLNE 60
Qy	61	TTKYLIGDGP	PEFGRCTQ	ELLFGK	SALIND	KARTAT	QPGTGCAL
Db	61	TTKYLIGDGP	PEFGRCTQ	ELLFGK	SALIND	KARTAT	QPGTGCAL
Qy	121	KRVVSNP	SWPNHKS	VFN	SAGLEV	REYAY	DAENHTL
Db	121	KRVVSNP	SWPNHKS	VFN	SAGLEV	REYAY	DAENHTL
Qy	181	CHNPTGIDPT	LEQWOT	LQALS	VEK	GWLP	LPFA
Db	181	CHNPTGIDPT	LEQWOT	LQALS	VEK	GWLP	LPFA
Qy	241	ASSYSK	QFGLN	RVG	ACT	LV	AA
Db	241	ASSYSK	QFGLN	RVG	ACT	LV	AA
Qy	301	DALRAINE	QELT	DM	RQIR	QRM	QLF
Db	301	DALRAINE	QELT	DM	RQIR	QRM	QLF

Db 301 DALRAIWQELTDMRQIRQMRQLFVNTLQEKGNRDFSLIKQNGMFSFSLGTLKEQVLR 360
Qy 361 LREEFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 396
|||||
Db 361 LREEFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 396

RESULT 2

US-10-673-786A-2
; Sequence 2, Application US/10673786A
; Publication No. US20040132165A1
; GENERAL INFORMATION:
; APPLICANT: AKHVERDIAN, VALERY ZAVENOVICH
; APPLICANT: SAVRASOVA, EKATERINA ALEKSEEVNA
; APPLICANT: KAPLAN, ALLA MARKOVNA
; APPLICANT: LOBANOV, ANDREY OLEGOVICH
; APPLICANT: KOZLOV, YURI IVANOVICH
; TITLE OF INVENTION: METHOD FOR PRODUCING L-THREONINE USING BACTERIA
; TITLE OF INVENTION: BELONGING TO THE GENUS ESCHERICHIA
; FILE REFERENCE: US-115
; CURRENT APPLICATION NUMBER: US/10/673,786A
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: PCT/JPO3/02067
; PRIOR FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: 2002104983
; PRIOR FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-673-786A-2

Query Match 100.0%; Score 2045; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 4.2e-190;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETGTPTVLTSVKKAQVYLLNE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETGTPTVLTSVKKAQVYLLNE 60
Qy 61 TTKNYLGDIGIPEFGRCCTOELLFGKGSALINDKARTATQPGTGCALRVAADFLAKNTSV 120
Db 61 TTKNYLGDIGIPEFGRCCTOELLFGKGSALINDKARTATQPGTGCALRVAADFLAKNTSV 120
Qy 121 KRVVSNPSWPNHKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
Db 121 KRVVSNPSWPNHKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFAYQGFARGLEEDAAGLRAFAAMHKELIV 240
Db 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFAYQGFARGLEEDAAGLRAFAAMHKELIV 240
Qy 241 ASSYSKNFGLNRYGVGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVVTATILSN 300
Db 241 ASSYSKNFGLNRYGVGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVVTATILSN 300
Qy 301 DALRAIWQELTDMRQIRQMRQLFVNTLQEKGNRDFSLIKQNGMFSFSLGTLKEQVLR 360
Db 301 DALRAIWQELTDMRQIRQMRQLFVNTLQEKGNRDFSLIKQNGMFSFSLGTLKEQVLR 360
Qy 361 LREEFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 396
Db 361 LREEFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 396

RESULT 3

US-10-369-493-21125
; Sequence 21125, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21125
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-21125

Query Match 80.9%; Score 1654; DB 4; Length 396;

Best Local Similarity 78.3%; Pred. No. 5.2e-152;
Matches 310; Conservative 34; Mismatches 52; Indels 0; Gaps 0;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETGTPTVLTSVKKAQVYLLNE 60
Db 1 MFEKITAAPADPILGLADLFRADERPGKINLIGVYKDETGTPTVLTSVKKAQVYLLNE 60
Qy 61 TTKNYLGDIGIPEFGRCCTOELLFGKGSALINDKARTATQPGTGCALRVAADFLAKNTSV 120
Db 61 TTKNYLGDIGIPEFGRCCTOELLFGKGSALINDKARTATQPGTGCALRVAADFLAKNTSV 120
Qy 121 KRVVSNPSWPNHKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
Db 121 KRVVSNPSWPNHKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFAYQGFARGLEEDAAGLRAFAAMHKELIV 240
Db 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFAYQGFARGLEEDAAGLRIFTKNHNLIV 240
Qy 241 ASSYSKNFGLNRYGVGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVVTATILSN 300
Db 241 ASSYSKNFGLNRYGVGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVVTATILSN 300
Qy 301 DALRAIWQELTDMRQIRQMRQLFVNTLQEKGNRDFSLIKQNGMFSFSLGTLKEQVLR 360
Db 301 DEFKAEMIQELATWRERIRMRQLFVNTLQEKGNRDFSLIKQNGMFSFSLGTLKEQVDR 360
Qy 361 LREEFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 396
Db 361 LREEFGVYAVSSGRINVAGLTLENMVPPLCEAIVAVL 396

RESULT 4

US-10-369-493-231
; Sequence 231, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 231
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus


```
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7497
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-369-493-7497

Query Match      48.7%; Score 996; DB 4; Length 396;
Best Local Similarity 48.2%; Pred. No. 6.8e-88;
Matches 191; Conservative 67; Mismatches 138; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETKTPTVLTSVKKAEOYLLENE 60
Db 1 LFSAVELAPRDPILGLNEAFNADTRTTKVNILGVGVYFNEEGKIPLLRAVRDAEKARVDA 60
Qy 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTAOPTGGTGALRVAADFLAKNTSV 120
Db 61 LPRGYLPIEGIAAYDAVQKLLGNDSPLIAAGRVVTAQAALGGTGALKIGADFLLKRLNPN 120
Qy 121 KRVVSNPSPNHNKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
Db 121 AKVAISDPSPWENHRALEFEGAGFEVVSYPYDAHTHGVNFDAMLNSALNSYAAGTIVVLHAC 180
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKWLPLDFPAYQGFARGLEEDAEGLRAFAAMHKLIV 240
Db 181 CHNPTGVDLNIQWQVVEVVKARNLVPFLDIAVQGFQNDIEADAAAVRLFAAAELNVFV 240
Qy 241 ASSYSKNGFLYNERVGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVVATILSN 300
Db 241 SSSFSKSFSLYGERVGALSIITASKEEAARVLSQKRVIRTNYSNPPTHGSSVVAALAS 300
Qy 301 DALRAIWEQELTDMRQIORMROLFVNTLOEKGANRDFSFIIKQNGMFSFSLTKQVLR 360
Db 301 PELRATWETELAEMDRIRAMRNGLVERLKASGVDRDFSFVNAQRGMFSYSLGTAPQVDR 360
Qy 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREEFGIYAVSTGRICVAALNTRNLDAVANAIAHVL 398

RESULT 9
US-10-369-493-8424
; Sequence 8424, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8424
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Ralstonia metallidurans
US-10-369-493-8424

Query Match      48.0%; Score 982; DB 4; Length 397;
Best Local Similarity 48.6%; Pred. No. 1.6e-86;
Matches 192; Conservative 61; Mismatches 142; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETKTPTVLTSVKKAEOYLLENE 60
Db 3 LFSAVEMAPRDPILGLNEAFNADTRASKVNLGVGVYTFDEGKIPLLRAVQAEKARLUTA 62
Qy 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTAOPTGGTGALRVAADFLAKNTSV 120
Db 63 TPRGYLPIEGIAAYDQAVQILLFKESPLITEGRVVTAAALGGTGALKIGADFLLKRLYPN 122
```

QY 121 KRVSNSPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLEAQAQGVVLFHGC 180
 D 122 ASVAISDPSWENHRAALFEAAGFPVYVYAYDAASHGLNFAGMVESLKSYPANTIVVLHAC 182
 QY 181 CHNPTGIDPTLEQWOTLAQLSVEKQWLPFLDFPAYOGFARGLEDEAGLRAPFAAMHKEIV 240
 D 182 CHNPTGIDPTLEQWOTLAQLSVEKQWLPFLDFPAYOGFARGLEDEAGLRAPFAAMHKEIV 240
 QY 241 ASSYSKNFGLYNERVAGCTLVAAADSETVDRAFSQMAAIRANYSNPPAHGASVVAATILSN 300
 D 242 ASSYSKNFGLYNERVAGCTLVAAADSETVDRAFSQMAAIRANYSNPPAHGASVVAATILSN 300
 QY 301 DALRAIWEQELTDMQRQIORMQLFVNTLOEKGANRDFSFIIKQNGMFSFSGLTKEQVLR 360
 D 302 DALRAIWEQELTDMQRQIORMQLFVNTLOEKGANRDFSFIIKQNGMFSFSGLTKEQVLR 360
 QY 361 LREBFGVYAVASGRVNVAGMTPDNMAPLCEAIVAV 395
 D 362 LREBFGVYAVASGRVNVAGMTPDNMAPLCEAIVAV 395
 RESULT 10
 US-10-369-493-15828
 ; Sequence 15828, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 15828
 ; LENGTH: 397
 ; TYPE: PRT
 ; ORGANISM: Xanthomonas campestris
 US-10-369-493-15828
 Query Match 47.7%; Score 975; DB 4; Length 397;
 Best Local Similarity 48.9%; Pred. No. 7.6e-86;
 Matches 193; Conservative 68; Mismatches 132; Indels 2; Gaps 2;
 QY 2 FENITAAPADPILGLADLFRADERPGKINLGIVYKDETGTPTVLTSSVKKAEQVLLNET 61
 D 4 FANVEQVPGDPIGLTEAYNADSRPNKVLGVGYDENGRIPLLRVAVQKIEQQALDAK 63
 QY 62 TKNYLGDIGPEFGRCQTELLFGKGSALINDKARTAQTCGTGALRVAADFLAKNTSVK 121
 D 64 PRGYLPIDGLAAYDKATQELLFGAESALVAGSRVATSTVCGSGALRVGADLLKLLPTS 123
 QY 122 RVVNSPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLEAQAQGVVLFHGC 181
 D 124 TIAISNPSWENHRAVFGAAGFEVVDYTFDAASHGLNFDGMLADLAKLEPGTVVLLHACC 183
 QY 182 HNPFGIDPTLEQWOTLAQLSVEKQWLPFLDFPAYOGFARGLEDEAGLRAPFAAMHKE-LIV 240
 D 184 HNPFGIDPTLEQWOTLAQLSVEKQWLPFLDFPAYOGFARGLEDEAGLRAPFAAMHKE-LIV 240
 QY 241 ASSYSKNFGLYNERVAGCTLVAAADSETVDRAFSQMAAIRANYSNPPAHGASVVAATILSN 300
 D 242 ASSYSKNFGLYNERVAGCTLVAAADSETVDRAFSQMAAIRANYSNPPAHGASVVAATILSN 300
 QY 301 DALRAIWEQELTDMQRQIORMQLFVNTLOEKGANRDFSFIIKQNGMFSFSGLTKEQVLR 360
 D 302 DALRAIWEQELTDMQRQIORMQLFVNTLOEKGANRDFSFIIKQNGMFSFSGLTKEQVLR 360
 QY 361 LREBFGVYAVASGRVNVAGMTPDNMAPLCEAIVAV 395
 D 362 LREBFGVYAVASGRVNVAGMTPDNMAPLCEAIVAV 395
 RESULT 12
 US-10-369-493-15460
 ; Sequence 15460, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng

QY 361 LREBFGVYAVASGRVNVAGMTPDNMAPLCEAIVAV 395
 D 362 LREBFGVYAVASGRVNVAGMTPDNMAPLCEAIVAV 395
 RESULT 11
 US-10-369-493-16208
 ; Sequence 16208, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 16208
 ; LENGTH: 397
 ; TYPE: PRT
 ; ORGANISM: Xanthomonas campestris
 US-10-369-493-16208
 Query Match 47.7%; Score 975; DB 4; Length 397;
 Best Local Similarity 48.9%; Pred. No. 7.6e-86;
 Matches 193; Conservative 68; Mismatches 132; Indels 2; Gaps 2;
 QY 2 FENITAAPADPILGLADLFRADERPGKINLGIVYKDETGTPTVLTSSVKKAEQVLLNET 61
 D 4 FANVEQVPGDPIGLTEAYNADSRPNKVLGVGYDENGRIPLLRVAVQKIEQQALDAK 63
 QY 62 TKNYLGDIGPEFGRCQTELLFGKGSALINDKARTAQTCGTGALRVAADFLAKNTSVK 121
 D 64 PRGYLPIDGLAAYDKATQELLFGAESALVAGSRVATSTVCGSGALRVGADLLKLLPTS 123
 QY 122 RVVNSPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLEAQAQGVVLFHGC 181
 D 124 TIAISNPSWENHRAVFGAAGFEVVDYTFDAASHGLNFDGMLADLAKLEPGTVVLLHACC 183
 QY 182 HNPFGIDPTLEQWOTLAQLSVEKQWLPFLDFPAYOGFARGLEDEAGLRAPFAAMHKE-LIV 240
 D 184 HNPFGIDPTLEQWOTLAQLSVEKQWLPFLDFPAYOGFARGLEDEAGLRAPFAAMHKE-LIV 240
 QY 241 ASSYSKNFGLYNERVAGCTLVAAADSETVDRAFSQMAAIRANYSNPPAHGASVVAATILSN 300
 D 242 ASSYSKNFGLYNERVAGCTLVAAADSETVDRAFSQMAAIRANYSNPPAHGASVVAATILSN 300
 QY 301 DALRAIWEQELTDMQRQIORMQLFVNTLOEKGANRDFSFIIKQNGMFSFSGLTKEQVLR 360
 D 302 DALRAIWEQELTDMQRQIORMQLFVNTLOEKGANRDFSFIIKQNGMFSFSGLTKEQVLR 360
 QY 361 LREBFGVYAVASGRVNVAGMTPDNMAPLCEAIVAV 395
 D 362 LREBFGVYAVASGRVNVAGMTPDNMAPLCEAIVAV 395
 RESULT 12
 US-10-369-493-15460
 ; Sequence 15460, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15460
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-15460

Query Match 47.7%; Score 975; DB 4; Length 398;
Best Local Similarity 48.9%; Pred. No. 7.6e-86;
Matches 193; Conservative 68; Mismatches 133; Indels 2; Gaps 2;
Qy 2 FENITAAPADPILGLADLFRADERPGKINLIGIVYKDTGKTPVLTSVKKAEOYLLENET 61
Db 2 FANVEQVPGDPILGLTEAYNADSRPNKVNGLGVGIYDENGRIPLLRVAVQKIEOQLALDAK 61
Qy 62 TKNYLIGDIGIEFGRCCTQELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSVK 121
Db 62 PRGYLPIDGLAAYDKATQELLFGAESALVAVSGRVATSTQVGGSGALRVGADLLKKLLPTS 121
Qy 122 RWMVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDPDALINSINEAQAGDVVLFHGCC 181
Db 122 TIATSNPSWENHRAVFGAGGEVDVYTFDAASHCLNFDGMLADLAKLEPGTVVILLHACC 181
Qy 182 HNPTGIDPTLEQWQTLAQLSVEKGMWLPDFPAYQGFARGLEDAEGLRAFAMHKE-LIIV 240
Db 182 HNPTGADLSREQWQTLAQLSVEKGMWLPDFPAYQGFARGLEDAEGLRAFAMHKE-LIIV 240
Qy 241 ASSYSKNFGLYNERVAGCTLVAADETVDFRAFQSKAAIRANYSNPPAHGASVVATILSN 300
Db 242 ASSYSKSFSLYGERVAGCTLVAADETVDFRAFQSKAAIRANYSNPPAHGASVVATILSN 301
Qy 301 DALRAIWEQELTDMRQIRQRMQLFVNTLOEKGANRDFSIKQNGMFSFSLGTLKEQVLR 360
Db 302 PELRDVWEQELTDMRQIRQRMQLFVNTLOEKGANRDFSIKQNGMFSFSLGTLKEQVLR 360
Qy 361 LREEFGVYAVASGRVNVAGTMDNNMAPLCEAIVAV 395
Db 361 LRDEFAIVAGTGRICVAALSQRNLDYVAQAVATV 395

RESULT 13
US-10-369-493-14025
; Sequence 14025, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14025
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Pseudomonas fluorescens
US-10-369-493-14025

Query Match 46.3%; Score 947; DB 4; Length 398;
Best Local Similarity 46.5%; Pred. No. 4.1e-83;

Matches 184; Conservative 66; Mismatches 146; Indels 0; Gaps 0;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINLIGIVYKDTGKTPVLTSVKKAEOYLLENE 60
Db 3 LFSAVEMAPRDPILGLNEAFNADTRTNKVNGLGVYVYCNEEGRIPLLRAVIEAETIRAAQH 62
Qy 61 TTKNYLIGDIGIEFGRCCTQELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120
Db 63 ASRGYLPIDGIAAYDAQVQKLLFGNSPLLAAGRVITTAQVGGTGALKIGADFLKQLLPN 122
Qy 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDPDALINSINEAQAGDVVLFHGCG 180
Db 123 AVVAISDPSPWENHRALEFETAGFPVQNYRYDYDAATHDVNRAGLLEDLNLALPNSIVVLHAC 182
Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKGMWLPDFPAYQGFARGLEDAEGLRAFAMHKE-LIIV 240
Db 183 CHNPTGVDLSPADWNNVLEAVKAKGHVPFLDMAYQGFQGDGIAEDAAAVRLFASGLTFFV 242
Qy 241 ASSYSKNFGLYNERVAGCTLVAADETVDFRAFQSKAAIRANYSNPPAHGASVVATILSN 300
Db 243 SSSFSKSFSLYGERVAGALSIISSKEESARVLQVKKRVITNYSNPPHGHASIVAAVINS 302
Qy 301 DALRAIWEQELTDMRQIRQRMQLFVNTLOEKGANRDFSIKQNGMFSFSLGTLKEQVLR 360
Db 303 PELRAQWEAEALAEKRLIRGMRQMVALLAKKAPQDFSFVGRQGRGMFSYGLTTEQVHR 362
Qy 361 LREEFGVYAVASGRVNVAGTMDNNMAPLCEAIVAVL 396
Db 363 LRNEFGIYALDTGRICVAALNQSNIIEAVTDAIVQVI 398

RESULT 14
US-10-369-493-9458
; Sequence 9458, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9458
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Xylella fastidiosa
US-10-369-493-9458

Query Match 45.2%; Score 924; DB 4; Length 395;
Best Local Similarity 46.2%; Pred. No. 7e-81;
Matches 183; Conservative 77; Mismatches 134; Indels 2; Gaps 2;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINLIGIVYKDTGKTPVLTSVKKAEOYLLENE 60
Db 1 LFTDVELVPGDPILSLNDAYNADTRTNKVNGLGIYDESCICIFLLRAVQVVEQLAKHP 60
Qy 61 TTKNYLIGDIGIEFGRCCTQELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120
Db 61 KPRGYLPIDGLPAYIKATQQLLFQVDSILLTAGRVATSTQVGGSGALRVAAEVLKQVLPH 120
Qy 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDPDALINSINEAQAGDVVLFHGCG 180
Db 121 ATVAISRPSPWENHRALEFETAGFPVQNYRYDYDAATHDVNRAGLLEDLNLALPNSIVVLHAC 180
Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKGMWLPDFPAYQGFARGLEDAEGLRAFAMH-KELI 239

Job time : 168 secs

```
Db 181 CHNPTGADLSRDQWQKQVALFOERQLPCIDLAYQGFNOGIDADAYAIRLLAEGISNVV 240
QY 240 VASSYSKNFGLYNERVAGCTLVADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILS 299
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 241 VANSYSKSFSLYGERVGLSIVASNTEQAQAIQSQVKRIIRTIYSSPPAHGAYLVAGVLN 300
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 300 NDALRAIWEQELTDMRQIRQRLFVNTLQEKGNRDPFSFIKQNGMFSFSGLTKEQVL 359
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 301 SHELRTLWEQELTQMRIRHGLRAGLVARLKALGA-PEPDFIQRQAGMFSYSGLSKIQVD 359
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 360 RLREFFGVYAVASGRVNVAGTTPDNMAPLCEAIVAV 395
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 360 RLREFFGIYAVSSGRICVAALSQHKLEYVAQAVKV 395
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 15
US-10-369-493-9222
; Sequence 9222, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9222
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Xylella fastidiosa
US-10-369-493-9222

Query Match 44.8%; Score 917; DB 4; Length 395;
Best Local Similarity 45.5%; Pred. No. 3.4e-80;
Matches 180; Conservative 79; Mismatches 135; Indels 2; Gaps 2;

QY 1 MFENITAAPADPILGLADLFRADERPGKINIGVYKDETGTGTPVLTSVKKAEOYLLENE 60
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 LFTDVELVPGPILSLNDAYNADTRTNKVNIGIYCDSEGCIPLLRAVQQVEEQAKHP 60
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 61 TTKNYLGIDGIPFGRCQTELLFGKGSALINDKRARTATPGTGTALRVAADF LAKNTSV 120
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 61 KPRGYLPIDGLPAYIKATQQLLFGVDSLTLTAGRVATSTVGGSGALRVAAEVLKQVLPH 120
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 121 KRUVSNBSPNHNKSVFNSAGLEVREYAYYDAENHTLDPDALINSNEAQAGDVVLFHGC 180
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 121 ATVAISRPSENHRAFLTAAGFKIEDTYFDTLGHAVDFTGMVADLAKLPKTVVLLHGC 180
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGWMLPLFPAYOCFARGLEEDAEGLRFAAAMH- KELL 239
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 181 CHNPTGADLSRDQWQKQVALFOERQLPCIDLAYQGFNOGIDADAYAIRLLAEGISNVV 240
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 240 VASSYSKNFGLYNERVAGCTLVADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILS 299
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 241 VANSYSKSFSLYGERVGLSIVASNTEQAQAIQSQVKRIIRTIYSSPPAHGAYLVASVLN 300
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 300 NDALRAIWEQELTDMRQIRQRLFVNTLQEKGNRDPFSFIKQNGMFSFSGLTKEQVL 359
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 301 SHELRTLWEQELTQMRIRHGLRAGLVARLKALGV-PEPDFIQRQAGMFSYSGLSKIQVD 359
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 360 RLREFFGVYAVASGRVNVAGTTPDNMAPLCEAIVAV 395
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 360 RLREFFGIYAVSSGRICVAALSQHKLEYVAQAVKV 395
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 18:12:15 ; Search time 22 Seconds
(without alignments)
501.028 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045

Sequence: 1 MFENITAAPADPILGLADLF.....VAGMTDPNMAPLCEAIVAVL 396

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 161667 seqs, 27834885 residues

Total number of hits satisfying chosen parameters: 161667

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications_AA New:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2011	98.3	396	7	US-11-114-922-76
2	1317	64.4	423	6	US-10-467-657-6536
3	842	41.2	397	6	US-10-467-657-1060
4	831.5	40.7	397	6	US-10-979-821-32
5	831.5	40.7	397	7	US-11-114-922-32
6	818	40.0	419	6	US-10-979-821-6
7	818	40.0	419	7	US-11-114-922-6
8	815	39.9	389	6	US-10-979-821-2
9	815	39.9	389	7	US-11-114-922-2
10	810	39.6	419	6	US-10-979-821-4
11	810	39.6	419	7	US-11-114-922-4
12	794	38.8	430	6	US-10-878-556A-146
13	756	37.0	412	6	US-10-878-556A-149
14	699	34.2	447	7	US-11-245-400-17
15	632.5	30.9	412	6	US-10-979-821-8
16	632.5	30.9	412	7	US-11-114-922-8
17	153.5	7.5	429	6	US-10-793-626-3174
18	123.5	6.0	369	7	US-11-055-822-116
19	123.5	6.0	369	7	US-11-055-822-486
20	113	5.5	473	7	US-11-087-099-10299
21	111	5.4	452	6	US-10-467-657-1850
22	105.5	5.2	392	7	US-11-087-099-941
23	105.5	5.2	397	7	US-11-232-405A-42
24	102	5.0	473	7	US-11-087-099-1989
25	100.5	4.9	871	7	US-11-087-099-7516

26	99	4.8	393	6	US-10-979-821-10	Sequence 10, Appl
27	99	4.8	393	7	US-11-114-922-10	Sequence 10, Appl
28	98	4.8	399	7	US-11-096-568A-22461	Sequence 22461, A
29	98	4.8	432	7	US-11-096-568A-22460	Sequence 22460, A
30	96	4.7	434	7	US-11-096-568A-20937	Sequence 20937, A
31	96	4.7	462	7	US-11-096-568A-20936	Sequence 20936, A
32	96	4.7	505	7	US-11-096-568A-20935	Sequence 20935, A
33	96	4.7	623	7	US-11-072-512-2547	Sequence 2547, Ap
34	96	4.7	657	6	US-10-873-528-57	Sequence 57, Appl
35	95	4.6	389	6	US-10-467-657-4468	Sequence 4468, Ap
36	95	4.6	461	7	US-11-096-568A-10640	Sequence 10640, A
37	94.5	4.6	444	7	US-11-205-109-7	Sequence 7, Appli
38	94	4.6	395	7	US-11-087-099-8148	Sequence 8148, Ap
39	94	4.6	894	6	US-10-485-517-416	Sequence 416, App
40	93	4.5	395	7	US-11-087-099-11586	Sequence 11586, A
41	90	4.4	326	7	US-11-096-568A-22462	Sequence 22462, A
42	90	4.4	356	7	US-11-087-099-1503	Sequence 1503, Ap
43	89.5	4.4	394	6	US-10-510-386-208	Sequence 208, App
44	89.5	4.4	633	7	US-11-087-099-10840	Sequence 10840, A
45	88.5	4.3	395	6	US-10-467-657-3516	Sequence 3516, Ap

ALIGNMENTS

RESULT 1

US-11-114-922-76

; Sequence 76, Application US/11114922

; Publication No. US20050282260A1

; GENERAL INFORMATION:

; APPLICANT: MCFARLAN, PAULA M.

; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE

; TITLE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS

; FILE REFERENCE: 023829-0396

; CURRENT APPLICATION NUMBER: US/11/114,922

; CURRENT FILING DATE: 2005-04-26

; PRIOR APPLICATION NUMBER: 10/422,366

; PRIOR FILING DATE: 2003-04-23

; PRIOR APPLICATION NUMBER: 60/374,831

; PRIOR FILING DATE: 2002-04-23

; NUMBER OF SEQ ID NOS: 91

; SOFTWARE: PatentIn Ver. 3.3

; SEQ ID NO 76

; LENGTH: 396

; TYPE: PRT

; ORGANISM: Escherichia coli

; US-11-114-922-76

Query Match 98.3%; Score 2011; DB 7; Length 396;
Best Local Similarity 98.5%; Pred. No. 1.1e-161;
Matches 390; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADPRPGKINIGIGYKDETKTPTVLTSVKKAQYLLNE 60

Db 1 MFENITAAPADPILGLADLFRADPRPGKINIGIGYKDETKTPTVLTSVKKAQYLLNE 60

Qy 61 TTKYLVGIDGTPGRCCTQELLFGKSGALINDKARTATQTCGTCALRVAADFLAKNTSV 120

Db 61 TTKYLVGIDGTPGRCCTQELLFGKSGALINDKARTATQTCGTCALRVAADFLAKNTSV 120

Qy 121 KRWVSNPSWPNHKSFNAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180

Db 121 KRWVSNPSWPNHKSFNAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180

Qy 181 CHNFTGIDPTLEQWOTLAQLSVEKGLPLDFPAYQGFARGLEEDAEGLRAFAAMHKEIV 240

Db 181 CHNFTGIDPTLEQWOTLAQLSVEKGLPLDFPAYQGFARGLEEDAEGLRAFAAMHKEIV 240

Qy 241 ASSYSKNFGLYNERVGACTILVAADSETVDRAFSONKAAIRANYSNPPAHGASVVAATILSN 300

Db 241 ASSYSKNFGLYNERVGACTILVAADSETVDRAFSONKAAIRANYSNPPAHGASVVAATILSN 300

Qy 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQKGNRDFSFIIKQNGMFSFSLTKQVLR 360
Db 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQKGNRDFSFIIKQNGMFSFSLTKQVLR 360
Qy 361 LREBFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREBFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 2

US-10-467-657-6536
; Sequence 6536, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6536
; LENGTH: 423
; TYPE: PRI
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6536

Query Match 64.4%; Score 1317; DB 6; Length 423;
Best Local Similarity 61.9%; Pred. No. 3.3e-103;
Matches 244; Conservative 62; Mismatches 88; Indels 0; Gaps 0;

Qy 2 FENITAAPADPILGLADLFRADERPGKINLGIVYKDTGKTPVLTSVKKAEQVLLNET 61
Db 30 FKHTEAAPADPILGLGEAFKAETRPKEKVNGLGIVYKDSAGATPIVKAKEAEKRLLESET 89
Qy 62 TKNYLIGIDGPEFGRCCTQELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSVK 121
Db 90 TKNYLTIDGVADYNEQTOILLFGKDHEIYASRRAKTAQSLGTTGALRIAABFAKRLNAQ 149
Qy 122 RVVWSNPSWPNHKSFNVSAGLEVREYAYYDAENHTLDFDALINSNEAQAGDVVLFHGCC 181
Db 150 TIWISNTPWPNHNAITAKAVGIQDKFYRYIDAKHCLDWDGWIEDLNQAKGDIVLLHGCC 209
Qy 182 HNPTGIDPTLEQWOTLAQLSVEKGWLPFLDFAYQGFARGLEEDAEGLRAPAAHMKELIVA 241
Db 210 HNPTGIDPTPEQWETLAKLSAEKGWLPFLDFAYQGFNGLEEDAYGLRVFLKHNTLLIA 269
Qy 242 SSYSKNFGLYNERVGACTIONVAADSETVDRAFSQKAAIRANYSNPPAHGASVWATILSND 301
Db 270 SSYSKNFGYNERVGAFTVAEDDEETAARAHQSQIKTIIRTLYSNPASHGANTIALVLKND 329
Qy 302 ALRAIWEQELTDMRQRIORMRQLFVNTLQKGNRDFSFIIKQNGMFSFSLTKQVLR 361
Db 330 DUKAQWIAELDEMRGRIKAMQKQFVELLKAKGSTQDFDPIIEQNGMFSFSLTPEQVRL 389
Qy 362 REBFGVYAVASGRVNVAGMTPDNMAPLCEAIVAV 395
Db 390 KNEFAIYAVRSGRINNAVIGITDNDLYCESIVK 423

RESULT 3

US-10-467-657-1060
; Sequence 1060, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita

; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1060
; LENGTH: 397
; TYPE: PRI
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1060

Query Match 41.2%; Score 842; DB 6; Length 397;
Best Local Similarity 41.7%; Pred. No. 2.9e-63;
Matches 165; Conservative 68; Mismatches 163; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINLGIVYKDTGKTPVLTSVKKAEQVLLNE 60
Db 1 MYRHIEVYVPGDPILSLVETFKNDPRPEKVNLSIGIYFDDGMRMPVLESVSRAETARAAP 60
Qy 61 TKNYLIGIDGPEFGRCCTQELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120
Db 61 APSPLPMEGLDVIYRSVAVQHLLFGKGNPALAQGRIVTVTLGGSGALKVGADFJHRWPE 120
Qy 121 KRVVWSNPSWPNHKSFNVSAGLEVREYAYYDAENHTLDFDALINSNEAQAGDVVLFHGC 180
Db 121 ARAYVSODPTWDNHRGIFEGAGFEVGTYPYDPAIVGVKFDGMTAFFNTLPENSVLILHPC 180
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPFLDFAYQGFARGLEEDAEGLRAPAAHMKELIV 240
Db 181 CHNPTGVDMSERQWDEVLIQIKTRKLPFMDIAVQGGDLSDAYAVRKAVERMDLPLFV 240
Qy 241 ASSYSKNFGLYNERVGACTIONVAADSETVDRAFSQKAAIRANYSNPPAHGASVWATILSN 300
Db 241 SNSFSKNLSYGERVGGVSVCPNKEADLVFGQLKFTVRRIYSSPPAHGAYIAADVNS 300
Qy 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQKGNRDFSFIIKQNGMFSFSLTKQVLR 360
Db 301 SELVALQWNEVYMRDRIRAMRQKLYGVLTARIPDRDFTYFIKORGFMFGYGLSVGVRR 360
Qy 361 LREBFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LRDEFAYVLLDSGRMCMVAGLNTSNITYVADALAEVL 396

RESULT 4

US-10-979-821-32
; Sequence 32, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; APPLICANT: WEINER, DAVID P.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; FILE REFERENCE: 023829-0390
; CURRENT APPLICATION NUMBER: US/10/979,821
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77

[illegible]

Qy 241 ASSYSKNFLYNERVGACTLV--AADSETVDRAFQSMKAAIRANYSNPPAHGASVATIL 298
Db 263 AASCSCNFGIYRERTGILIAIGEAAGRTVQ---ANLNFARQNYSPFPDGHARLVMTIL 319
Qy 299 SNDALRAIWEOLTDMRQIORMQLFVNTLQ-EKGANRDFSFIKQNGMFSFSGLTKEQ 357
Db 320 EDETLASDWKAELEEVRLNMLTLRQLADALQAETGSNR-FGFVAEHRGMSRLGITPAE 378
Qy 358 VLRLEBEFGYVAVASGRVNVAGMTDPDNMAPLCEAIVAVL 396
Db 379 VERLRTHEGVYMGVDSRLNIAGLNRTTVPVLARAVAKVL 417
RESULT 7
US-11-114-922-6
; Sequence 6, Application US/11114922
; Publication No. US20050282260A1
; GENERAL INFORMATION:
; APPLICANT: HICKS, PAULA M.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; CURRENT FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Patentin Ver. 3.3
; SEQ ID NO 6
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Rhodobacter sphaeroides
US-11-114-922-6

Query Match 40.0%; Score 818; DB 7; Length 419;
Best Local Similarity 42.9%; Pred. No. 3.3e-61;
Matches 171; Conservative 61; Mismatches 157; Indels 10; Gaps 5;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINIGIGVYKDETGTPTVLTSVKKAQVLLNE 60
Db 26 MLTALKPQPADKILQIQMFREDARADKIDLGVGYYKDTGLTPVMRAVKAERKLEWE 85
Qy 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKRAARTQTPGGTGALRVAADFLAKNTSV 120
Db 86 TTKTYTGLAGEPAYNANAKLIL---AGAVPADRVASVATPGGTGAVRQALELRMASPE 142
Qy 121 KRVMVSNPSPNHNKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
Db 143 ATVWISNPTWPNHLSIVKYLIGIPMEYRYFDAETCAVDAEGLMEDLAQVKAGDVVLLHGC 202
Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKWLPLDFPAYOGFARGLEEDAEGLRAFAAMHKLIV 240
Db 203 CHNPTGANPNVQWLAVCESLARTGAVPLIDLAYQGGDGLMDAAATRLATRLPEVLI 262
Qy 241 ASSYSKNFLYNERVGACTLV--AADSETVDRAFQSMKAAIRANYSNPPAHGASVATIL 298
Db 263 AASCSCNFGIYRERTGILIAIGEAAGRTVQ---ANLNFARQNYSPFPDGHARLVMTIL 319
Qy 299 SNDALRAIWEOLTDMRQIORMQLFVNTLQ-EKGANRDFSFIKQNGMFSFSGLTKEQ 357
Db 320 EDETLASDWKAELEEVRLNMLTLRQLADALQAETGSNR-FGFVAEHRGMSRLGITPAE 378
Qy 358 VLRLEBEFGYVAVASGRVNVAGMTDPDNMAPLCEAIVAVL 396
Db 379 VERLRTHEGVYMGVDSRLNIAGLNRTTVPVLARAVAKVL 417
RESULT 8
US-10-979-821-2

; Sequence 2, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; APPLICANT: WEINER, DAVID P.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0390
; CURRENT APPLICATION NUMBER: US/10/979,821
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patentin Ver. 3.3
; SEQ ID NO 2
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Sinorhizobium meliloti
US-10-979-821-2
Query Match 39.9%; Score 815; DB 6; Length 389;
Best Local Similarity 44.3%; Pred. No. 5.4e-61;
Matches 170; Conservative 63; Mismatches 137; Indels 14; Gaps 4;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINIGIGVYKDETGTPTVLTSVKKAQVLLNE 60
Db 1 MFEDALARQADDPDLLFLGLFKDPERPGKVLGVGYRDETGTPTIFRAVKAERKLETTQ 60
Qy 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKRAARTQTPGGTGALRVAADFLAKNTSV 120
Db 61 DSKAYTGPEDLVPLDLRLWELV---GDTTIERSHVAGVQTFGGSGALRLAADLIAR-MGG 116
Qy 121 KRVMVSNPSPNHNKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
Db 117 RGIWGLSPNPNHAPIFKAAGLDIATYDFDIPSQSVIFDNLVSALEGAASGDVALLHAS 176
Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKWLPLDFPAYOGFARGLEEDAEGLRAFAAMHKLIV 240
Db 177 CHNPTGVLSEAOQWMEIALVAERGLLPLVDLAYQGGFGRGLDQDVAGLRHLLGVVPEALV 236
Qy 241 ASSYSKNFLYNERVGACTLVAAADSETVDRAFQSMKAAIRANYSNPPAHGASVATILSN 300
Db 237 AVSCSKSFGLYRERAGAI PARTSTASADRVSRNLGARTSYSNMPPDHGAUVRTILDD 296
Qy 301 DALRAIWEOLTDMRQIORMQLFVNTL---QEKGANRDFSFIKQNGMFSFSGLTKE 356
Db 297 PELARDTEELTEMLRLMTGLRRSLAEGRLTRWQSLGAVAD-----QEGMFSMLPLSEA 350
Qy 357 QVLRLEBEFGYVAVASGRVNVAGM 380
Db 351 EVMLRTEHGIYMPASGRINIAGL 374
RESULT 9
US-11-114-922-2
; Sequence 2, Application US/11114922
; Publication No. US20050282260A1
; GENERAL INFORMATION:
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; CURRENT FILING DATE: 2005-04-26

;; PRIOR APPLICATION NUMBER: 10/422,366
;; PRIOR FILING DATE: 2003-04-23
;; PRIOR APPLICATION NUMBER: 60/374,831
;; PRIOR FILING DATE: 2002-04-23
;; NUMBER OF SEQ ID NOS: 91
;; SOFTWARE: PatentIn Ver. 3.3
;; SEQ ID NO 2
;; LENGTH: 389
;; TYPE: PRT
;; ORGANISM: Sinorhizobium meliloti
US-11-114-922-2

Query Match 39.9%; Score 815; DB 7; Length 389;
Best Local Similarity 44.3%; Pred. No. 5.4e-61;
Matches 170; Conservative 63; Mismatches 137; Indels 14; Gaps 4;

QY 1 MFENITAAPADPILGLADLFRADERPGKINIGVYKDETKGTPTVLTSVKKAEOYLLENE 60
Db 1 MFDALARQADDPLFLIGLFRKDERPGKVDLGVGVYRDETGRTPIPRAVKAEEKRLLEQ 60
QY 61 TTKNYLGIDGIPFGRCCTQELLFGKGSALINDKRAARTAOCTPGTGALRVAADFLAKNTSV 120
Db 61 DSKATIGPSEDVFLDLRLWELV--GGDTIERSHVAGVOTPGSGALRLAADLIAR-MGG 116
QY 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSINEAQAGDVVLFHGC 180
Db 117 RGIMGLPSPNHNAPIFKAAGLDIATYDFDIPSQSVIFDNLVSALEGAASGDVALLHAS 176
QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPFLFPAYQGFARGLEEDAEGLRAPAMHKLIV 240
Db 177 CHNPTGGVLSAQMEIALVAERGLLPLVDLAYQGFGRGLDQVAGLRHLLGVVPEALV 236
QY 241 ASSYSKNFLGNYRERVAGTCTVADSETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
Db 237 AVSCSKFGLYRERAGATPARTSSTASADRVSNLAGLARTSYSMPPDHGAAVRTILD 296
QY 301 DALRAIWEQELTDMRQIORMQLFVNTL-----QEGKANRDFSFIKQGMFSGLTKE 356
Db 297 PELRRDWTETMLRMTGLRSLAELRLTRWQSLGAVAD-----QEGMFSMLPLSEA 350
QY 357 QVLRREEFGVYAVASGRVNVAGM 380
Db 351 EVMRLRTEGHYMPASGRINIAGL 374

RESULT 10
US-10-979-821-4
; Sequence 4, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; APPLICANT: WEINER, DAVID P.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; FILE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0390
; CURRENT APPLICATION NUMBER: US/10/979,821
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 4
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Rhodobacter sphaeroides

US-10-979-821-4
Query Match 39.6%; Score 810; DB 6; Length 419;
Best Local Similarity 42.4%; Pred. No. 1.6e-60;
Matches 169; Conservative 62; Mismatches 158; Indels 10; Gaps 5;

QY 1 MFENITAAPADPILGLADLFRADERPGKINIGVYKDETKGTPTVLTSVKKAEOYLLENE 60
Db 26 MLTALKPOPADKILQLIQMFREDARADKIDLGVGYYKDPGTGLTPMRAVKAEEKRLWEVE 85
QY 61 TTKNYLGIDGIPFGRCCTQELLFGKGSALINDKRAARTAOCTPGTGALRVAADFLAKNTSV 120
Db 86 TTKTYTGLADEPAYNAAMAKLIL---AGAVPADRVASVATPGTGAVRQALELIRNASPE 142
QY 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSINEAQAGDVVLFHGC 180
Db 143 ATWISNTWPNHLSIVKYLGPMEYRYFYDAETGAVDAEGNMEDLAQVKGADVVLLHGC 202
QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPFLFPAYQGFARGLEEDAEGLRAPAMHKLIV 240
Db 203 CHNPTGANPNPVQWLAIACESLARTGAVPLIDLAYQGFQGLEMDAAATRLATRLPEVLI 262
QY 241 ASSYSKNFLGNYRERVAGTCTV--AADSETVDRAFQSKAAIRANYSNPPAHGASVATIL 298
Db 263 AASCSKNFGIYRERTGILIAIGEAAGRTVQ---ANLNFNLNRQNTSYFPPDHGARLVMTIL 319
QY 299 SNDALRAIWEQELTDMRQIORMQLFVNTLQ-EKGANRDFSFIKQGMFSGLTKEQ 357
Db 320 EDETLASDWKALEBVLRLNMLTLRQLADALQAGTGSNR-FGFVAEHRGMSRLGITPAE 378
QY 358 VLRREEFGVYAVASGRVNVAGMTPDNNMAPLCEAIVAVL 396
Db 379 VERLRTEGHYMPVGVSDRLNIAGLNRITTVVPLARAVAKVL 417

RESULT 11
US-11-114-922-4
; Sequence 4, Application US/11114922
; Publication No. US20050282260A1
; GENERAL INFORMATION:
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; FILE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; CURRENT FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 4
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Rhodobacter sphaeroides
US-11-114-922-4

Query Match 39.6%; Score 810; DB 7; Length 419;
Best Local Similarity 42.4%; Pred. No. 1.6e-60;
Matches 169; Conservative 62; Mismatches 158; Indels 10; Gaps 5;

QY 1 MFENITAAPADPILGLADLFRADERPGKINIGVYKDETKGTPTVLTSVKKAEOYLLENE 60
Db 26 MLTALKPOPADKILQLIQMFREDARADKIDLGVGYYKDPGTGLTPMRAVKAEEKRLWEVE 85
QY 61 TTKNYLGIDGIPFGRCCTQELLFGKGSALINDKRAARTAOCTPGTGALRVAADFLAKNTSV 120
Db 86 TTKTYTGLADEPAYNAAMAKLIL---AGAVPADRVASVATPGTGAVRQALELIRNASPE 142
QY 121 KRVVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSINEAQAGDVVLFHGC 180

Db 143 ATWISNTPWPHLSIVKVLGTPMREYRYFDAETCAVDAEGMEDLAQVKAQGVVLLHGC 202
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFFAVQGFARGLEDEAEGURAFAMHKEIV 240
Db 203 CHNPTGANPNVQWLAICESLARTGAVPLIDLAVQGFQDGLMDAAATRLLATRLPEVLI 262
Qy 241 ASSYSKNFGLYNERVGACTLV--AADSETVDRAFSQMKAAIRANYSNPPAHGASVVATIL 298
Db 263 AASCSKNFGIVRETTGIIAIEGAAGRTVQ---ANLNFLLARQNYSPFPDGHARLVWTIL 319
Qy 299 SNDALRAIWEOLTDMRORIORMROLFVNTLQ-EKGANRDFSFIKQNGMFSFSGLTKEQ 357
Db 320 EDETLASDKAELEVRNLMLTLRQLADALQAEATGSNR-FGFVAEHKGMFSRLGIPTAE 378
Qy 358 VLRLEEFVYAVASGRVNVAGMTPDNMAPICEAIVAVL 396
Db 379 VERLRTGVMGVGDSRLNIAGLNRTTVPVLARAVAKVL 417

RESULT 12

US-10-878-556A-146
; Sequence 146, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 146
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: sw_hum/aatm_human
; DATABASE ENTRY DATE: 1986-07-21
US-10-878-556A-146

Query Match 38.8%; Score 794; DB 6; Length 430;

Best Local Similarity 41.2%; Pred. No. 3.6e-59;
Matches 161; Conservative 67; Mismatches 161; Indels 2; Gaps 2;

Qy 4 NITAAPADPILGLADLFRADERPGKINLIGVYKDETGTPTVLTSVKKAEOYLLENETTK 63
Db 35 HVEMGPPDPILGVTEAFKEDTNSKKNLGVGAYRDDNGKPYVLPVSRKAEAQIAAKNLDK 94
Qy 64 NYLGIDGIDPEFCRGTQELLFGKGSALINDKRAKRTAQTGGTGALRVAADFLAKNTSVKR- 122
Db 95 EYLPDGLAEFCFAELALGENSEVLKSGREVTVTQTTSGTGALRIGASFQRPFKFSRD 154
Qy 123 VVSVSNPWNHKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAQGVVLPFGCCH 182
Db 155 VFLPKPTWGNHTPTFRDAGMQLQGGRYDYDPKCGDFGTGAVDEDSIKIPEQSVLLHLHAC 214
Qy 183 NPTGIDPTLEQWOTLAQLSVEKGWLPDFFAVQGFARG-LEBDAEGLRAFAAMHKEIVA 241
Db 215 NPTGVDPRPEQNKETATVVKKRNLAFFADMAVQGFASGDDGDADAVRHFIEQGINVCLC 274
Qy 242 SSYSKNFGLYNERVGACTLVAAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSD 301
Db 275 QSYAKMGLYGERVGAFTVCKDAEAKRVESQLKILIRPMYSNPPLNGARITAAAILNTP 334
Qy 302 ALRAIWEOLTDMRORIORMROLFVNTLQEKANRDFSFIKQNGMFSFSGLTKEQVLR 361
Db 335 DUKQWLQEVKGMDRIIGRTQLVSNLKGESTHWHQHTDQIGMFCETGLKPKQVERL 394
Qy 362 RBEFGVYAVASGRVNVAGMTPDNMAPLCEAI 392
Db 395 IKFESIYMTKDRISVAGVTSNVGYLAHI 425

RESULT 13

US-10-878-556A-149
; Sequence 149, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 149
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: sw_hum/aatc_human
; DATABASE ENTRY DATE: 1990-08-01
US-10-878-556A-149

Query Match 37.0%; Score 756; DB 6; Length 412;

Best Local Similarity 43.0%; Pred. No. 5.4e-56;
Matches 166; Conservative 54; Mismatches 156; Indels 10; Gaps 5;

Qy 10 ADPIL--GLADLFRADERPGKINLIGVYKDETGTPTVLTSVKKAEOYLL-ENETTKNYL 66
Db 12 AQPVLVFKLTADFPDDPRKRVGLGVAYRTDDCHPWLPVVKVQEQXIANDNSLNHEYL 71
Qy 67 GIDGIDPEFCRGTQELLFGKGSALINDKRAKRTAQTGGTGALRVAADFLAK-----NTSVK 121
Db 72 PILGLAEFSCASRLALGDDSPALKEXRGVQSLGGTGALRIGADFLARWYNGTNNKVT 131
Qy 122 RVVSVSNPWNHKSVFNSAGL-EVREYAYDAENHTLDFDALINSLNEAQAQGVVLPFGC 180
Db 132 PVYVSSPTWENHNAVFSAGFKDIRSYRYMDAERKGLDQGLFLNDLENAPESFVWLHAC 191
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFFAVQGFARG-LEBDAEGLRAFAAMHKELI 239
Db 192 AHNPTGIDPTPEQWKQIASVNMKGRFLPFPFDSATQGFASGNLERDANRAIRFVSEGEFF 251
Qy 240 VASSYSKNFGLYNERVGACTLVAAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILS 299
Db 252 CAQSFKNFGLYNERVGNLTIVGKEPESILQVLSQMEKIVRIITWSNPPAQCARIVASTLS 311
Qy 300 NDALRAIWEOLTDMRORIORMROLFVNTLQEKANRDFSFIKQNGMFSFSGLTKEQV 359
Db 312 NPFLFEETWGNVKTMDRILTMRSRLARLEALKTPGTWNHITDQIGMFSFTGLNPKQVE 371
Qy 360 RLREFGVYAVASGRVNVAGMTPDNM 385
Db 372 YLVNEKHIYLLPSGRINVSGLTTKNL 397

RESULT 14

US-11-245-400-17
; Sequence 17, Application US/11245400
; Publication No. US20060040357A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhkar
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Encoding Melanoma
; TITLE OF INVENTION: Associated Antigen Molecules, Aminotransferase
; TITLE OF INVENTION: Molecules, ATPase Molecules, Acyltransferase Molecules,
; TITLE OF INVENTION: Pyridoxal-Phosphate Dependant Enzyme Molecules and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 35800/247400
; CURRENT APPLICATION NUMBER: US/11/245,400
; CURRENT FILING DATE: 2005-10-06
; PRIOR APPLICATION NUMBER: US/10/164,966
; PRIOR FILING DATE: 2002-06-07

```

; PRIOR APPLICATION NUMBER: 10/034,864
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/258,517
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 09/996,194
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/250,348
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,073
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/253,878
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,338
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 09/908,928
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/220,465
; PRIOR FILING DATE: 2000-07-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam Aminotransferase family model
US-11-245-400-17

Query Match      34.2%; Score 699; DB 7; Length 447;
Best Local Similarity 41.4%; Pred. No. 3.8e-51;
Matches 182; Conservative 60; Mismatches 152; Indels 46; Gaps 13;

Qy      1 MFENITAAPADPILGLADLFRADERPGKIN---LGIGVYKDETGTPTVLTSVKKAEOYL 56
Db      4 MAANVSHGPGDPIFGVAFKEDPRGKDNPNVIGVAYEPQLGKDLVLPVAKKAERKL 63
Qy      57 -----LENETTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTATQTPGGTGAALRV 109
Db      64 ALDREGNIEFREIKYELPIHGLPEFREAIKLLFGARSPKLFKEVRVVTQLGGTGAALRL 123
Qy      110 AADFLAK---NTSVKRVVNSPNSPNHK-SVFNISAGLEV-REYAYY-DAENHTLDFDALI 163
Db      124 AADFLANPGDSRGSRGVLVPTTPWNYKRDIFWAAAGVEVIVPYHYKDDENNFGLDFALE 183
Qy      164 NSLNEAQAGDV--VLFHGCCNPTGIDPTLEQWTLAQLSVEKGLPLDFEAYQGFARG 220
Db      184 AAIEKAPEKNTKTVLLHNNPHNPTGDTPTREQLKKIADVVKERNILLSLDEAYQGFVFG 243
Qy      221 -LEEDAEGLRAPAAHMK-----ELIVASSYSKNFGLYNERVGACTLV----- 261
Db      244 SLDEDAASVAEFAEVEKMECNGDELLVVSQFSKNFGLYGRVGAIVVNPRTGDAVIS 303
Qy      262 -AADSETVDRAFSPQMAAIRANYSN---PPAHGASVVAITILSDNALRAIWEQELTDMRQR 317
Db      304 AAAMSSAGRVSSQLQALARAMYNSPDPDPHGAIEIVARILERRDLFTSWLEEVKGMACR 363
Qy      318 IQMRV-QLFVNTLQKGNRDFSFIIKQNGHFSGLTKEQVLR--EEFGVAVASGR 374
Db      364 IPNGRLYLMMDLRLKLLKEEDDWSHIIIEQEGMFSFTWLLNEEQVNVSPGSEFHIYEPGWGR 423
Qy      375 VNVAGMTPDNNAPLCEAIVA 394
Db      424 ISLAGLSEANVEEAERIRA 443

```

```

RESULT 15
US-10-979-821-8
; Sequence 8, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.

```

```

; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; APPLICANT: WEINER, DAVID P.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; FILE REFERENCE: 023829-0390
; CURRENT APPLICATION NUMBER: US/10/979,821
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 8
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Leishmania major
US-10-979-821-8

Query Match      30.9%; Score 632.5; DB 6; Length 412;
Best Local Similarity 38.0%; Pred. No. 1.3e-45;
Matches 150; Conservative 62; Mismatches 174; Indels 9; Gaps 6;

Qy      2 FENITAAPADPILGLADLFRADERPGKINLGIVYKDETGTPTVLTSVKKAEOYLLENET 61
Db      13 WQKIQAPADPVIDFLAKRAAAAGP-KANLVIGAYRDEQGRPYPLRVVRKABQLLLDMNL 71
Qy      62 TKNYLGIDGIPFGRCTQELLFGKGSALINDKARTATQTPGGTGAALRVAAADFLAK-NTS 119
Db      72 DYEYLPISGYQPFIDEAVKIIYGNVLEEN---LVAVOTLSGTGAVSLGAKLLTRVFDAL 128
Qy      120 VKRVVNSPNSPNHKSVNSAGLE-VREYAYYDAENHTLDFDALINSLENAQAQDVVLPH 178
Db      129 TPIYLSPTWPNHYGVVKAAGWKNICTYAYYDPKTVSLNFEKMKDILAAPDGSVFFILH 188
Qy      179 GCCHNPTGIDPTLEQWTLAQLSVEKGLPLDFEAYQGFARG-LEEDAEGLRAPAAHMK 237
Db      189 QCAHNPTGVDPSEQMNEIASLMLAKHQVFFDSAYQGYASGSLDTPDAYAARLFAARRGIE 248
Qy      238 LIVASSYSKNFGLYNERVGACTLVAADSETVDRAFSPQMAAIRANYSNPPAHGASVVAITI 297
Db      249 VLLAQSFKNWGLYSERAGTSLLLKDKTKRADVKSVWDSLIREEYTCPPAHGARLAHLI 308
Qy      298 LNSDALRAIWEQELTDMRQRIQMRQLFVNTLQKGNRDFSFIIKQNGHFSGLTKREQ 357
Db      309 LSNNELKKEWEAEELSAMAERIRTMRTVYDELLRLQTPGSEHVINQIGMFSFLGLSKAQ 368
Qy      358 VLRLREFGVYAVASGRVNVAGMTPDNNAPLCEAI 392
Db      369 -CEYCONHNIPITVSGRANWAGLTHTETALMLAQTI 402

Search completed: March 14, 2006, 18:15:20
Job time : 24 secs

```

This Page Blank (uspio)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2006, 14:05:42 ; Search time 5402 Seconds
(without alignments)
10315.330 Million cell updates/sec

Title: US-10-673-786A-1

Perfect score: 1191

Sequence: 1 atgtttgagacattaccgc.....cgattgtgcagtgctgtaa 1191

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	855	71.8	864	10	CL660432
2	783.4	65.8	785	10	CL695845
3	563	47.3	671	10	CL695172
4	556.8	46.8	645	6	CA093007
5	482.4	40.5	557	6	CF306417
6	445.8	37.4	468	1	AW221110
7	236.8	19.9	1690	4	AX103658
8	228.8	19.2	1749	8	AX109239
9	224.8	18.9	1832	4	AX105250
10	220.4	18.5	2452	4	CR857622
11	215	18.1	516	6	CF338891
12	211.6	17.8	2447	4	CR860094
13	205.4	17.2	1599	4	CNS0A2EC
14	203	17.0	977	7	CK259742
15	200.6	16.8	1486	4	CNS092WQ
16	199.6	16.8	902	7	CK259741
17	190.4	16.0	1643	4	CNS0A80C
18	190	16.0	858	7	CO101379
19	189.6	15.9	919	7	CO732049
20	188.2	15.8	968	7	CK262759
21	188	15.8	704	3	BJ624279
22	187	15.7	914	7	CO732002

23	185.6	15.6	852	7	CO459606	CO459606	MZCCL2002
24	185.6	15.6	876	8	DR059430	DR059430	RTNITI 17
25	184.2	15.5	759	6	CD919780	CD919780	G608.114J
26	183	15.4	854	7	CO464147	CO464147	MZCCL1504
27	182.6	15.3	783	8	DN147062	DN147062	4853 A11
28	182.2	15.3	1564	4	CNS0EPVU	CR636228	Tetraodon
29	179.8	15.1	847	6	CD574536	CD574536	UCRPT01_0
30	179.4	15.1	2017	4	CR860603	Pongo pyg	
31	179	15.0	1131	7	CV734832	Cr12 11 H	
32	179	15.0	1440	4	CNS0SOL3	CR634545	Tetraodon
33	177.8	14.9	1496	4	CR616132	CR616132	full-length
34	177.8	14.9	1509	4	CR628608	CR628608	full-length
35	177.8	14.9	1512	4	CR605482	CR605482	full-length
36	177.8	14.9	1525	4	CR615083	CR615083	full-length
37	177.8	14.9	1540	4	CR598904	CR598904	full-length
38	177.8	14.9	1560	4	CR622320	CR622320	full-length
39	177.8	14.9	1563	4	CR601723	CR601723	full-length
40	177.8	14.9	1719	4	CR606661	CR606661	full-length
41	177.8	14.9	1959	4	CR625122	CR625122	full-length
42	177.8	14.9	1979	4	CR619243	CR619243	full-length
43	177.8	14.9	1979	4	CR620721	CR620721	full-length
44	177.6	14.9	1030	9	BZ567307	BZ567307	paca2-164
45	176.6	14.8	1517	4	CNS0GMY6	CR725768	Tetraodon

ALIGNMENTS

RESULT 1
CL660432
LOCUS
DEFINITION
PRI0137a.D10 - PRI0137a.B21 (864) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.
ACCSSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
COMMENT

CL660432 864 bp DNA linear GSS 09-JUL-2004
PRI0137a.D10 - PRI0137a.B21 (864) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.
CL660432 GI:50145720
GSS.
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 864)
Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
AppADB: an AcedB database for the nematode satellite organism Pristionchus pacificus
Nucleic Acids Res. 32 (1), D421-D422 (2004)
14681447
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.
Location/Qualifiers
1. .864
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus var. California"
/note="Vector: pBifos-5 Fosmid vector"

FEATURES

source
1. .864
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus var. California"
/note="Vector: pBifos-5 Fosmid vector"

ORIGIN

Query Match 71.8%; Score 855; DB 10; Length 864;
Best Local Similarity 100.0%; Pred. No. 9.4e-257;
Matches 855; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 55 CTGTTTCGCGGACGTACGCTCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAGAT 114
Db 10 CTGTTTCGCGGACGTACGCTCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAGAT 69
Qy 115 GAGACGGCAAAACCCCGGTACTGACACGCGTGAAAAAGGCTGAACAGTATCTCTCGAA 174
Db 70 GAGACGGCAAAACCCCGGTACTGACACGCGTGAAAAAGGCTGAACAGTATCTCTCGAA 129
Qy 175 AATGAAACACCAAAATTAACCTCGGCAATGACGCGATCCCTGAATTTGGTGTCTGCACT 234
Db 130 AATGAAACACCAAAATTAACCTCGGCAATGACGCGATCCCTGAATTTGGTGTCTGCACT 189
Qy 235 CAGGAACCTGCTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTCGCACGCA 294
Db 190 CAGGAACCTGCTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTCGCACGCA 249
Qy 295 CAGACTCCGGGGGACCTGGCGCCTACGCGTGGCTGCGGATTTCTTGGCAAAAAATACC 354
Db 250 CAGACTCCGGGGGACCTGGCGCCTACGCGTGGCTGCGGATTTCTTGGCAAAAAATACC 309
Qy 355 AGCGTTAAGCTGTGGGTGAGCAACCCAAAGCTGGCCGAAACCATAGAGCGCTTTTAAC 414
Db 310 AGCGTTAAGCTGTGGGTGAGCAACCCAAAGCTGGCCGAAACCATAGAGCGCTTTTAAC 369
Qy 415 TCTGACGCTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAATCACACTCTGAC 474
Db 370 TCTGACGCTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAATCACACTCTGAC 429
Qy 475 TTCATGCACTGATTAAACAGCTGAAATGAAGCTCAGGCTGGCGAGCTAGTGTCTGTTCCAT 534
Db 430 TTCATGCACTGATTAAACAGCTGAAATGAAGCTCAGGCTGGCGAGCTAGTGTCTGTTCCAT 489
Qy 535 GCGTCTGCCATAACCCAAACCGGTATCGACCTACGCTTGGAAACAAATGGCAAAACACTGCA 594
Db 490 GCGTCTGCCATAACCCAAACCGGTATCGACCTACGCTTGGAAACAAATGGCAAAACACTGCA 549
Qy 595 CAACCTCCGTTGAGAAAGGCTGGTACGCTGTTTGAATCTCGCTTACAGGGTTTGGCC 654
Db 550 CAACCTCCGTTGAGAAAGGCTGGTACGCTGTTTGAATCTCGCTTACAGGGTTTGGCC 609
Qy 655 CCGTCTCTGGAAGAGATGCTGAAGGACTGGCGGCTTTTCGGGCTATGCATAAAGAGCTG 714
Db 610 CCGTCTCTGGAAGAGATGCTGAAGGACTGGCGGCTTTTCGGGCTATGCATAAAGAGCTG 669
Qy 715 ATTGTTCCAGTTCCTACTCTAAATAAATTTGGCCCTGTATACACGAGCGTGTGGCGCTGT 774
Db 670 ATTGTTCCAGTTCCTACTCTAAATAAATTTGGCCCTGTATACACGAGCGTGTGGCGCTGT 729
Qy 775 ACTCTGGTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGCG 834
Db 730 ACTCTGGTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGCG 789
Qy 835 ATTCCGCTTAACCTAAACCCACGACACACGCGCTTCTGTTGTTGCCACCATCTG 894
Db 790 ATTCCGCTTAACCTAAACCCACGACACACGCGCTTCTGTTGTTGCCACCATCTG 849
Qy 895 AGCAACGATGGCTTA 909
Db 850 AGCAACGATGGCTTA 864
```

```
RESULT 2
LOCUS CL695845
DEFINITION PR1017b H12.2 - PRI017b.BR (785) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
CL695845
VERSION CL695845.1 GI:50217753
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
```

```
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 785)
Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
AppADB: an Acedb database for the nematode satellite organism
Pristionchus pacificus
Nucleic Acids Res. 32 (1), D421-D422 (2004)
14681447
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.
FEATURES
Location/Qualifiers
source 1..785
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
/notes="Vector: pEpifos-5 Fosmid vector"
ORIGIN
Query Match 65.8%; Score 783.4; DB 10; Length 785;
Best Local Similarity 99.9%; Pred. No. 3e-234;
Matches 784; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 380 ACCCAAGCTGCGCGAACCATTAAGAGCGCTCTTTAACTCTGCAAGTCTGGAGTTTCGTGAAT 439
Db 1 ACCCAAGCTGCGCGAACCATTAAGAGCGCTCTTTAACTCTGCAAGTCTGGAGTTTCGTGAAT 60
Qy 440 ACGCTTATTATGATCGGAAATACACTCTTCGATTCGATGCACTGATTAAACAGCCTGA 499
Db 61 ACGCTTATTATGATCGGAAATACACTCTTCGATTCGATGCACTGATTAAACAGCCTGA 120
Qy 500 ATGAAGCTCAGGCTGGCGACGTAGTCTGTTCCATGGCTGTCGCATAACCCACCGGTA 559
Db 121 ATGAAGCTCAGGCTGGCGACGTAGTCTGTTCCATGGCTGTCGCATAACCCACCGGTA 180
Qy 560 TCGACCCCTACGCTGGAAACAATGGCAACAACCTCGGTTGAGAAAGCGTGGT 619
Db 181 TCGACCCCTACGCTGGAAACAATGGCAACAACCTCGGTTGAGAAAGCGTGGT 240
Qy 620 TACCGCTGTTTGACTTCGCTTACCAGGGTTTTCGCCGCTGCTGGAAGAGATGCTGAAG 679
Db 241 TACCGCTGTTTGACTTCGCTTACCAGGGTTTTCGCCGCTGCTGGAAGAGATGCTGAAG 300
Qy 680 GACTGCGCGCTTTTCGGGCTATGCATAAAGAGCTGATTGTTGCCAGTTCCTACTCTAAA 739
Db 301 GACTGCGCGCTTTTCGGGCTATGCATAAAGAGCTGATTGTTGCCAGTTCCTACTCTAAA 360
Qy 740 ACTTTGGCCTGTACAACGAGCGTGTGGCGCTTGTACTCTGTTGCTGCGACAGTGA 799
Db 361 ACTTTGGCCTGTACAACGAGCGTGTGGCGCTTGTACTCTGTTGCTGCGACAGTGA 420
Qy 800 CCGTTGATCGGCATTCAGCCAAATGAAGGGCGATTTCGGCTAACTACTCTAACCCAC 859
Db 421 CCGTTGATCGGCATTCAGCCAAATGAAGGGCGATTTCGGCTAACTACTCTAACCCAC 480
Qy 860 CAGCACACGCGCTTCTGTTGTTGCCACCATCTCGAGCAACGATGCGTGTAGTCGATTT 919
Db 481 CAGCACACGCGCTTCTGTTGTTGCCACCATCTCGAGCAACGATGCGTGTAGTCGATTT 540
Qy 920 GGGAAACAAGAGCTGACTGATATGCGCAGCGATTTTCAGCGTATGCGTCA 979
Db 541 GGGAAACAAGAGCTGACTGATATGCGCAGCGATTTTCAGCGTATGCGTCA 600
```

```
Qy 980 ATACGCTCAGGAAAGGCGCAACCGCGACTTCAGCTTTATCATCAAAAGAGACGCGCA 1039
Db 601 ATACGCTCAGGAAAGGCGCAACCGCGACTTCAGCTTTATCATCAAAAGAGACGCGCA 660
Qy 1040 TGTCTCCTTCAGTGGCGCTGACAAAGACAAAGTGTCTGCTGCGGCAAGATTGGCG 1099
Db 661 TGTCTCCTTCAGTGGCGCTGACAAAGACAAAGTGTCTGCTGCGGCAAGATTGGCG 720
Qy 1100 TATATGCGGTTCCTTCGTCGCGTAAATGTGGCGGGATGACACACAGATAACATGGCTC 1159
Db 721 TATATGCGGTTCCTTCGTCGCGTAAATGTGGCGGGATGACACACAGATAACATGGCTC 780
Qy 1160 CGCTG 1164
Db 781 CGCTG 785

RESULT 3
CL695172 671 bp DNA linear GSS 10-JUL-2004
LOCUS PRI0166a.G11.2 - PRI0166a.BR (671) Mixed stage fosmid library of P.
DEFINITION pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
ACCESSION CL695172
VERSION 1 GI:50217080
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE
AUTHORS Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
TITLE AppaDB: an AcedB database for the nematode satellite organism
JOURNAL Pristionchus pacificus
PUBMED Nucleic Acids Res. 32 (1), D421-D422 (2004)
COMMENT 14681447
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.
FEATURES
source
1. .671
Location/Qualifiers
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
/note="Vector: pBifos-5 Fosmid vector"

ORIGIN
Query Match 47.3%; Score 563; DB 10; Length 671;
Best Local Similarity 99.0%; Pred. No. 5.6e-165;
Matches 577; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 610 AAAGGCTGGTTACCGCTGTTTGACTTCGCTTACCAGGGTTTGGCCGCTGCTCGAGAA 669
Db 1 AAAGGCTGGTTACCGCTGTTTGACTTCGCTTACCAGGGTTTGGCCGCTGCTCGAGAA 60
Qy 670 GATGCTGAAGGACTCGCGCTTTTCGCGCTATGCTAAAGAGCTGATTGTGCCAGTTCC 729
Db 61 GATGCTGAAGGACTCGCGCTTTTCGCGCTATGCTAAAGAGCTGATTGTGCCAGTTCC 120
Qy 730 TACTCTAAAAATTTGGCCTGTACACAGAGCGTGTGGCGCTTCTACTCTGTGTTGCTGCC 789
Db 121 TACTCTAAAAATTTGGCCTGTACACAGAGCGTGTGGCGCTTCTACTCTGTGTTGCTGCC 180
```

```
Qy 790 GACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGGCGATTTCGGCTAACTAC 849
Db 181 GACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGGCGATTTCGGCTAACTAC 240
Qy 850 TCTAACCCACAGACACGCGCTTCCTGTTGTTGCCACCATCTCTGAGCAAGATGCGTTA 909
Db 241 TCTAACCCACAGACACGCGCTTCCTGTTGTTGCCACCATCTCTGAGCAAGATGCGTTA 300
Qy 910 CGTG- CGATTGTTGGGAAACAAGAGCTGACTGATATGCGGCAGCGTATTACAGCGTATGCGTCA 968
Db 301 CGGCCGATTGGGAAACAAGAGCTGACTGATATGCGGCAGCGTATTACAGCGTATGCGTCA 360
Qy 969 GTTGTTCGTCAATACGCTGCGAGGAAAGGCGCAAAACCGCGACTTCAGCTTTATCATCAA 1028
Db 361 GTTGTTCGTCAATACGCTGCGAGGAAAGGCGCAAAACCGCGACTTCAGCTTTATCATCAA 420
Qy 1029 ACAGAACGGCATGTTCTCTTCAGTGGCTGACAAAAGAAACAAGTGTGCTGCTCGCGGA 1088
Db 421 ACAGAACGGCATGTTCTCTTCAGTGGCTGACAAAAGAAACAAGTGTGCTGCTCGCGGA 480
Qy 1089 AGAGTTTGGCGTATATCGGTTGCTTCTGTCGCGTAAATGTCGCCGGATGACACCAGA 1148
Db 481 AGAGTTTGGCGTATATCGGTTGCTTCTGTCGCGTAAATGTCGCCGGATGACACCAGA 540
Qy 1149 TAACATGGCTCCGCTGTGCGAAGCGATTGTGCGAGTGTCTGTAA 1191
Db 541 TAACATGGTCCGCTGTGCGAAGCGATTGTGCGAGTGTCTGTAA 583

RESULT 4
CA093007 645 bp mRNA linear EST 23-SEP-2003
LOCUS SCCCL1002F08.b CL1 Saccharum officinarum cDNA clone SCCCL1002F08
DEFINITION 3', mRNA sequence.
ACCESSION CA093007
VERSION CA093007.1 GI:34946314
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE
1 (bases 1 to 645)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCSEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: 002 row: F column: 08
Seq primer: M13/Forward primer.
FEATURES
source
1. .645
Location/Qualifiers
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCCCL1002F08"
/lab_host="XLIBlue MRF"
/clone_lib="CL1"
/note="Organ: Pool of sugarcane calli submitted to low
(40C) and high (37 C) temperature stress; Vector:
pBluescript; Site 1: EcoRI; Site 2: XhoI; An
unidirectional cDNA library generated from (pool of
sugarcane calli submitted to low (40C) and high (37 C)
```



```
VERSION AW221110.1 GI:6532794
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 468)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., and
Rinning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D., and
Giovannoni,J.
TITLE Generation of ESTs from tomato fruit tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
source Location/Qualifiers
1..468
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clones="cLEF12H11"
/tissue_type="fruit pericarp"
/dev_stages="mature green (3-5 days pre-ripening)"
/lab_host="SOUR"
/clone_lib="tomato fruit mature green, TAMU"
/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLEP - Fruit were tagged at the 1cm stage and
harvested 3-5 days prior to ripening. Fruit were cut in
half to verify the seeds were indeed 'immature' and the
seeds and locules were discarded prior to freezing the
pericarp"
ORIGIN
Query Match 37.4%; Score 445.8; DB 1; Length 468;
Best Local Similarity 99.6%; Pred. No. 3.5e-128;
Matches 447; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 153 GCCTGAACAGTATCTGCTCGAAATGAAACCAACCAAAATTTACCTCGGCATTGACGGCAT 212
Db 1 GCCTGAACAGTATCTGCTCGAAATGAAACCAACCAAAATTTACCTCGGCATTGACGGCAT 60
Qy 213 CCCTGAATTTGGTGGCTGCACTCAGGAATCTGCTGTTTGGTAAAGTAGCGCCCTGATCAA 272
Db 61 CCCTGAATTTGGTGGCTGCACTCAGGAATCTGCTGTTTGGTAAAGTAGCGCCCTGATCAA 120
Qy 273 TGACAAAGCTGCTCGACGGCACAGACTCGGGGGGCACTGGCGCACTACCGGTGGGTGC 332
Db 121 TGACAAAGCTGCTCGACGGCACAGACTCGGGGGGCACTGGCGCACTACCGGTGGGTGC 180
Qy 333 CGATTTCTCGCAAAATACACAGCTTAAGCGTGTGGGTGAGCAACCAAGCTGCC 392
Db 181 CGATTTCTCGCAAAATACACAGCTTAAGCGTGTGGGTGAGCAACCAAGCTGCC 240
Qy 393 GAACCATAGAGCGCTTTAACTCTGCAGGCTCTGGAAGTTCGTGAATACGCTTATTATGA 452
Db 241 GAACCATAGAGCGCTTTAACTCTGCAGGCTCTGGAAGTTCGTGAATACGCTTATTATGA 300
Qy 453 TGGCGAAAATCACACTCTTGACTTCGATGCACTGATTAACAGCCTGAATGAAGCTCAGGC 512
Db 301 TGGCGAAAATCACACTCTTGACTTCGATGCACTGATTAACAGCCTGAATGAAGCTCAGGC 360
Qy 513 TGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 572
Db 361 TGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Qy 573 GGAACAATGGCAACACTGGCACAACTCT 601
421 GGAACAATGGCAACACTGGCACAACTAT 449
RESULT 7
AY103658 1690 bp mRNA linear HTC 18-FEB-2005
LOCUS Zea mays P00118382 mRNA sequence.
DEFINITION Zea mays P00118382 mRNA sequence.
ACCESSION AY103658
VERSION AY103658.1 GI:21206736
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1690)
AUTHORS Gardiner,J., Schroeder,S., Polacco,M.L., Sanchez-Villeda,H.,
Fang,Z., Morgante,M., Landewe,T., Fengler,K., Useche,F.,
Hanafey,M., Tingey,S., Chou,H., Wing,R., Soderlund,C. and Coe,E.H.
Jr.
TITLE Anchoring 9,371 maize expressed sequence tagged unigenes to the
bacterial artificial chromosome contig map by two-dimensional
overgo hybridization
JOURNAL Plant Physiol. 134 (4), 1317-1326 (2004)
PUBMED 15020742
REFERENCE 2 (bases 1 to 1690)
AUTHORS Haney,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 3 (bases 1 to 1690)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
source Location/Qualifiers
1..1690
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeGDB:636895"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
ORIGIN
Query Match 19.9%; Score 236.8; DB 4; Length 1690;
Best Local Similarity 51.5%; Pred. No. 2.9e-62;
Matches 597; Conservative 0; Mismatches 557; Indels 6; Gaps 2;
Qy 22 GTCCTCGCGACCCGATTCGGGCTCGCGCATCTGTTTCGTCGGCATGACGTCGCCGC 81
Db 264 GCCCGGAGGACCCCATCTCGGAGTGACCGTTGCCCTACACAAAGGATCCCGAGCCCGTG 323
Qy 82 AAAATTAACTTCGGGATTGGTCTATAAGATGAGACGGCGCAAAACCCCGTACTGACC 141
Db 324 AAGTCAACTTCGGGGTTCGGCGCCTACCGGACCGGAGGAGGAGCCCTAGTCTGAAC 383
Qy 142 AGCGTGAATAAGGCTGAACAGTATCTCTCGAAATGAACACAC- - -CAAAATTTACCTC 198
```

Db	384	GTGGTCAGCGCGCGGACCAAAATGTTGATCAATAATCCGTACGTCACGTCACGAGTACCTA	443
Qy	199	GGCAATTGACGGCATCCCTGGAATTTGGTCGCTGCACTACGGAACCTGCTGTTTGGTAAAGGT	258
Db	444	CCAATCACCGGCTGCTGGCTGGAATCAATAAGCTGAGCGCTAAGCTTATCTTTGGCGCTGAC	503
Qy	259	AGCGCCCTGATCAATGACAAACGCTCTGCGACGGCACAGACTCCGGGGGACCTGGCGCA	318
Db	504	AGCCCTGCTATTCAGGAGAAATAGGCTTCTACCGTGCAGTCCCTATCGGGTACTGGTTCT	563
Qy	319	CTAGCGTGGCTGCGATTTCTTCCGCAAAAATACACGCGTTAAGCGTGTGGGTGAGC	378
Db	564	TTAAGAGTCGAGGTGATTTCTTTCGAAGCACTATCACGAGCGCACTATCTACATCCCA	623
Qy	379	AACCCAAGCTCGCGCAACCATAGAGCGCTTTTAACTCTGAGGTCTGGAAAGTTCGTGAA	438
Db	624	CAACCAACCTGGGAAATCACCCAAAAGCTTTCACCCCTATCTGGCTTGAACGTTAGGAGC	683
Qy	439	TACGCTTATTATGTCGGAAATCACACTCTTGACTTCGATCTCGATGACTGATTAACAGCTG	498
Db	684	TACCGCTATTATGATCTCTGCAACATGCGAGCCTTCATTCGAAGCACTCTCGAAGACCTC	743
Qy	499	AATGAAGCTCAGGCTGGCGAGCTAGTGTCTGTTCATGCTGCTGCCTAATCAACCAACCGGT	558
Db	744	GGTCTGCTCTTCAAGTTCAATTTGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	803
Qy	559	ATCGACCTACGCTGGAAACATAGCAACATGCGACAACTCTCTCGTTGAGAAAGGTGG	618
Db	804	GTAGATCTTACATCGAACAGTGGGAACAGATTAGGCAGCTGATGAGATCAAAATCACTG	863
Qy	619	TTACCGCTGTTGACTTCGCTTACCAGGCTTTTGGCGTG---GTCGTGAAGAGATGCT	675
Db	864	CTTCGGTCTTTGACAGTGCCTATCAAGGCTTTGCAAGTGGAAAGTCTTGACANAAGTCT	923
Qy	676	GAAGGACTGCGCGCTTTCGCGCTATGCATAAAGAGCTGATTTGTCAGTTCCTACTCT	735
Db	924	CAGTCAGTCGCTATGTTTGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	983
Qy	736	AAAACTTTGGCTGTACAAACGAGCGTGTGGCGCTGTACTCTCTGTTGTCGCGACAGT	795
Db	984	AAGACATGGGATTTGATGGAGAGCGTGTGGCGCTTTGAGCATTTGATGTAAGAGTGCC	1043
Qy	796	GAAACCGTTGATCGCGCATTCAGCCAAATGAAGCGGCGATTCGCGCTAATCTACTTAAC	855
Db	1044	GATGTAGCTGTAGGGTTGAAAGTCAACTCAAACTTGTCAATCAGGCGCTATGATTTCAAAC	1103
Qy	856	CCACCAGCACAGCGCGCTTCTGTTGTCGCCATCTCTGAGCAACGATGCGTTAGCTGCG	915
Db	1104	CCTCCTCTTATGTTGCTCTATCTGTTGCTACCATCTCAGGGACAGCGAGATGTTCAAC	1163
Qy	916	ATTTGGGAACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGTATGCGTCACTGTTTC	975
Db	1164	GAAATGACTCTGGAATCAAGGCCATGCTGATAGGATCAATTAACATGAGGCAACAATA	1223
Qy	976	GTCATAGCTGTCAGGAAAAGGCGCAACCGCGACTTCAGCTTTATCATCAACAGAAC	1035
Db	1224	TTTATAGCTGAAATCCAGAGGAACCCCTGTGTGATTTGAGCGCATATATTAAGCAAT	1283
Qy	1036	GGCATGTTCTCTTCAGTGGCTGACAAAAGCAAGTGTGCGTCTCGCGGAGAGTTT	1095
Db	1284	GGGATGTTTACTTTTCACTGGGCTGAAATAGCGAAACAAAGTCGCATTCATGAGGAGGATAC	1343
Qy	1096	GGCGTATATGCGGTTGCTTCTGTCGCTAAATGTGGCGGGAATGACACAGATAAATG	1155
Db	1344	CACATTTATGACATCTGATGGGAGGATCAGACTGCGCGGTTTGAGCATGAGGACTGTG	1403
Qy	1156	GCTCCGCTGTGGGAAGCGAT	1175
Db	1404	CCCATCTTGCAGATGCCAT	1423
RESULT 8			
CX109239			

LOCUS	CX109239		
DEFINITION	RECM0798 A normalized whole-life-cycle cDNA library of rice <i>Oryza sativa</i> (indica cultivar-group) cDNA clone B1130A13, E1078H06, B1100J15, B1107E19, B1131F09, B1075E21, B1075', mRNA sequence.		
ACCESSION	CX109239		
VERSION	CX109239.1		
KEYWORDS	EST.		
SOURCE	Oryza sativa (indica cultivar-group)		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		
REFERENCE	1 (bases 1 to 1749)		
AUTHORS	Zhang,J., Feng,Q., Jin,C., Qiu,D., Zhang,L., Xie,K., Yuan,D., Han,B., Zhang,Q. and Wang,S.		
TITLE	Features of the expressed sequences revealed by a large-scale analysis of ESTs from a normalized cDNA library of the elite indica rice cultivar Minghui 63		
JOURNAL	Plant J. 42 (5), 772-780 (2005)		
PUBMED	15918889		
COMMENT	Contact: Wang S National Key Laboratory of Crop Genetic Improvement Hauzhong Agricultural University Wuhan 430070, China Tel: 86-27-87282044 Fax: 86-27-87287092 Email: shiyingwang@hotmail.com		
FEATURES	Seq primer: T7.		
source	1. .1749		
	/organism="Oryza sativa (indica cultivar-group)"		
	/mol_type="mRNA"		
	/strain="indica"		
	/cultivar="Minghui 63"		
	/db_xref="taxon:39946"		
	/clone="B1130A13, E1078H06, B1100J15, B1107E19, B1131F09, B1075E21, B107"		
	/issue_type="whole plant"		
	/dev_stage="whole-life-cycle"		
	/lab_host="E. coli DH10B"		
	/clone_lib="A normalized whole-life-cycle cDNA library of rice"		
	/notes="vector: pSPORT1; Site 1: SalI; Site 2: NotI; The library is constructed based on the strategy of saturation hybridization with genomic DNA using rice cultivar Minghui 63. This library consists of cDNA from 15 directionally cloned cDNA libraries constructed with different tissues from 9 developmental stages."		
Query Match	19.2%;	Score 228.8;	DB 8; Length 1749;
Best Local Similarity	51.0%;	Pred. No. 9.6e-60;	
Matches	592;	Conservative 0;	Mismatches 562; Indels 6; Gaps 2;
Qy	22	GCTCTCTCCGACCGGATTTCTGGGCTCGCGCATCTGTTTCTGTCGCGGATGAACGTCCTCCGCGC	81
Db	302	CGCGCGGAGACCCCATCTCTCGGGTGACGCTCCGTCACACAGGATCCGAGCCCGCTG	361
Qy	82	AAATTAACCTCGGGATTTGTTGTTCTATAAAGATGAGACGGGCAAAACCCCGGTACTGACC	141
Db	362	AAGTCAATCTCGCGTCGCGGCTACCGGACCGGAGGAGGAGCCCTGCTGCTGAAT	421
Qy	142	AGCGTGAAGAGGTGAACAGATCTGCTCGAATAATGAACACCACCC---AAAAATTACCTC	198
Db	422	GTGTCAGGCGCGCTGAGCAGATGCTGATCAACACCCGTCACGTTTGAAGAGTATTG	481
Qy	199	GGCATTCAGCGCATCCCTGGAATTTGGTTCGCTCAGGAACCTGCTGTTTGTAAAGCT	258
Db	482	CCGATCACTGGACTGGCCGATTTCAATAAGCTGAGTCTAGCTTATTTTGGTCTGAC	541
Qy	259	AGCGCCCTGATCAATGACAAACGTCGTCGACCGGACAGACTCCGGGGGACCTGGCGCA	318
Db	542	AGTCTGCCATTCAAGAGATAGGTTGGCTACAGTTTCAAGTCTGTTGTCAGGAACCTGGTTCT	601

QY	319	CTAGCGTGGCTGCCGATTTCTCTGGCAAAAATACACAGGTTAAACGGTGTGTGGGTGAGC	378
QY	360	TTAAGGTTGGAGTGAATTTCTTGAAGGCATTATCATGAACGCATATATACATCCCA	661
QY	379	AACCAAGCTGGCGCAACATTAAGAGCGTCTTTAACTCTGCAGGCTCTGGAAGTTCTGTGA	438
QY	386	CAGCAACCTGGGGGAATCACCCAAAGTTTCACTTTAGCTGGCTGACCTGTAGAGT	721
QY	439	TACGCTTATATGATCGGAAATCACACTCTTGACTTCGATGCACTGATTAACAGCGCTG	498
QY	499	AATGAAGCTCAGGCTGGCAGTAGTCTGTTCCATGGCTGTGCGCATTAACCCAAACCGGT	558
QY	559	ATCGACCTACGCTGCAACAAATGGGCAACACACGTCGACCACTCTCGTTGGAAGGCTGG	618
QY	602	GTAGCCCAACTTTGACAGTGGGCAACAGATCAGGCACTGATGAGATCAAAAGCATG	901
QY	619	TTACCGCTGTTGACCTTCTACAGAGGTTTTGCGCGTG---GTCCTGGAAGAAGTCT	675
QY	676	GAAGGACTGGCGCTTTCGCGCTATGCAATAAGAGCTGATTTGTTCAGGTTCTTACTCT	735
QY	735	CAATCAGTGGCATGTTTGTCTGATGGTGGTGAATGCTCATGGCTCAGAGCTACGCT	1021
QY	796	AAACCTTTGGCTGTACAAAGCGGTGTGGCGCTTGTACTCTGTTGTCGCGCAGT	795
QY	855	AAGAAATGGGATTTGATGGAGAGCGGTGCGGTGCTTTAAGCATAGTTTGTGGAAGTCT	1081
QY	916	GAACACGCTGATCGCCATTCAGCCAAATGAAAGCGCGGATTTGCGCTTAACCTACTTAAC	855
QY	976	GTCAATACGCTGAGGAAAGCGCAACACCGGACTTTCAGCTTATCATCAACAGAAC	1035
QY	1036	GGCATGTTCTCTTACGTGGCTGCAAAAAGCAAGTGTCTGCGCTCTGCGGGAAGGTTT	1095
QY	1096	GGCGTATATGCGGTGCTTCTGCTGCGTGAATATGTCGCGGATGACACAGATAACATG	1155
QY	1156	GCTCCGCTGTGCAAGCGAT	1175
QY	1442	CCGCATCTTGCAGATGCAAT	1461
LOCUS	AY105250		
DEFINITION	Zea mays	PCO107699	mRNA sequence.
ACCESSION	AY105250		
VERSION	AY105250.1		GI:21208328
KEYWORDS	HTC.		
SOURCE	Zea mays		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD		

Db	580	CCATCTCGGGAAACCA	CACACCATCTT	CACGGATGCTGCATGCAAGCTTAC	AGGTTAT	639
Qy	442	GCTTATTATGATGCGGAAAT	CA	CACTCTTGACTTCGATGCACTGATTAAC	CAGCGCTGAAT	501
Db	640	CGGTATTATGACCCCA	GACTTGCGGTTT	TGACTTTCA	CAGGTGCTGGAGGATATT	699
Qy	502	GAAGCTCAGGCTGGCGA	CGTAGTCTGTCCATGGCTGCTGCCATAACCCAA	CGGTATC	561	
Db	700	AAAATACAGACGAGAGT	GTCTTCTTGTCATGCTGCGCCACAAT	CCCCACGGAGTG	759	
Qy	562	GACCTTAGCTGGAACAAT	TGGCAAACTCTGGCACA	ACTCTCCGTTGAGAAAGCGCTGTTA	621	
Db	760	GACCCGGTCCGGAACAG	TGGAGAAATAGCAACAGTGGTGAAGAAAGGAAT	CTCTTT	819	
Qy	622	CCGCTGTTTGA	CTTCACAGGGTTT	TGCCCGTGG--TCTGGAAGAGATGCTGAA	678	
Db	820	CGGTTCTTTGAC	TGCGCTTACCAAGCTT	TGCCAGTGTGATGAGATGCTCTGG	879	
Qy	679	GGACTGCGCGCTTT	CGCGGCTATGCATAAAGAGCTGATTTGTCGACGTCTCTACTCTAAA	738		
Db	880	GCTGTGCGCCACTT	CATCGAACAGGGCATTAATGTTTGTCTCTGCCAATCATATGCCAAG	939		
Qy	739	AATTTTGGCTGTACAA	CGAGCGTGTGGCGCTTGTACTCTGTGTGCTGCGACAGTGAA	798		
Db	940	AACATGGCTTATAT	TGTTAGCGGTGAGAGCCTTCACTATGGTCTGC	AAAGATGCGGAT	999	
Qy	799	ACGTTGATCGCGATTT	CAGCCAAATCAAGCGCGGATTCGCGTA	ACTACTTAACCCA	858	
Db	1000	GAAGCCAAAAGGGTAGAGT	CACAGTGAAGATCTTTGATCCGTC	CCATGTATTCACCCCT	1059	
Qy	859	CCAGCACACGCGCTT	CTGTGTTGGCCACATCTGAGCAACGATGCGTTACGTGCGATT	918		
Db	1060	CCCCTCAATGGGGCCGGAT	TGCTGCTGCCAATCTGAAACCCACAGATTTGCGNAACAA	1119		
Qy	919	TGGGAAACAAGAGCTG	ACTGATATCGCGCAGCGTATTCAGCGTATGCGTCA	AGTTGTTGCTC	978	
Db	1120	TGGCTGCAAGAGTGAAAGG	CATGGCCGACCGCATTTGCGATGCGGACTCAACTGGTC	1179		
Qy	979	AATACGCTGCAAGGAAAAG	CGGCAATTCAGCTTTATCATCAAAACAGACAGCGC	1038		
Db	1180	TCCAACTCAAGAAAGGAGG	TTCCACCCCAACTGGCAACACATCACCGCAAAATGGC	1239		
Qy	1039	ATGTTCTCTTTCAGTGGCCT	TGACAAAAGAACAGTGTGCGTCTGCGCAAGAGTTTGGC	1098		
Db	1240	ATGTTCTGTTTCA	CAGGCTTAAGCCTGAACAGGTGGAGCGCTGATCAAGGAGTTCTCC	1299		
Qy	1099	GTATATGCGGTTGCTT	CTGCTGCGGTAAATGTGGCCGGGATGA	CAACAGATAACATGGCT	1158	
Db	1300	ATCTACATGACA	AAAGATGCGCGCATCTCTGTGCGAGGGTCACTCCAGCAAGCTGGCG	1359		
Qy	1159	CCGCTGTGCGAAGCGATT	1176			
Db	1360	TACCTTGCCCATG	CCATT	1377		

RESULT 11	CF338891	516 bp	linear	EST 18-AUG-2003
LOCUS	CF338891/c			
DEFINITION	RCL1--03-B23.g1 Regenerated callus lambda phage cDNA library (RCL1) Oryza sativa (japonica cultivar-group) cDNA clone RCL1--03-B23, mRNA sequence.			
ACCESSION	CF338891			
VERSION	CF338891.1	GI:33826170		
KEYWORDS	EST.			
SOURCE	Oryza sativa (japonica cultivar-group)			
ORGANISM	Oryza sativa (japonica cultivar-group) Oryza sativa (japonica cultivar-group) Eukaryota: Viridiplantae: Streptophyta: Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.			
REFERENCE	1 (bases 1 to 516)			
AUTHORS	Kim, J. S., Jun, K. M., Cheong, P. J., Kim, M. J., Lee, T. H., Shin, Y. C., Song, S. I., Kim, J. K., Kim, Y.-K. and Nahm, B. H.			

FEATURES
source
 1. 516
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nackdong"
 /db_xref="taxon:39947"
 /clone="RCL1-03-B23"
 /tissue_type="callus"
 /dev_stage="proliferated callus on 2N6 media for 30 days"
 /lab_host="E.coli SOLR"
 /clone_lib="Regenerated callus lambda phage cDNA library (RCL1)"
 /notes="Vector: pBluescript SK(+); Site 1: SstI; Site 2: XhoI; cDNA was inserted into lambda Uni-ZAP XR vector at 5' end with SstI and 3' end with XhoI site. Callus was induced on 2N6 media for 30 days and cultured for 36hrs on regenerated media"

[illegible]

Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact R2PD for ordering:
<http://www.r2pd.de/cgi-bin/products/cl.cgi?ClonedID=DKFZp469B0125>
Further information about the clone and the sequencing project is
available at <http://mips.gsf.de/projects/cdna/>.

FEATURES

```

source
1..2447
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp469B0125"
/tissue_type="kidney"
/clone_lib="469 (synonym: pkidi). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stages="adult"
/notes="aspartate transaminase precursor, mitochondrial
(Homo sapiens)"
1..2447
/gene="DKFZp469B0125"
/join(103..1116,1116..1394)
/gene="DKFZp469B0125"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH92240.1"
/db_xref="GI:55731048"
/translations="KALLHSGRLPGIAAAPHPCLAARASARASWTHVEMGPPDI
LGVTEAKFRDINSKMLGVGYRDDNGKPYVPSVRKAEQIAAKNLDEYLPDGL
AEFCASAEALAGENSEVLKGRFVTVQTSIGTALRIGASFLQRFKSRDFLPKP
SWGHTTIFRDAGHOLQYRYDPKTFDGFVAVEDISKIPEQSVLLHACAHNPFG
VDPPEQWKEZTIVKRNLFPAFDMAYQGFASGDGDKAWAVRHFTFEGQINVCLOS
YAKNMGILGERVGAFTWVCKDAEKVESOLKILIRPMYSNPPLNGARAAATLNP
DLRTWLEQVKGMDRLIIGMRTOLVSLKKEGSHNHQIHITDQIGMFCFTGLKPEQVE
RLIKEFSIYMTKDGRIISVAGVISNVGLAHAIHQVTK"

```

ORIGIN

```

Query Match      17.8%; Score 211.6; DB 4; Length 2447;
Best Local Similarity 50.7%; Pred. No. 2.8e-54;
Matches 587; Conservative 0; Mismatches 564; Indels 7; Gaps 3;

QY 25  CTGCGCAGCCGATTCCTGGGCTGGCGATCTGTTTCGTCCGATGAACGTCCTGGCAAA 84
DB 220 CTTCCAGATCCATTCCTGGAGTCACCTGAAGCTTTAAGAGGACACACCAATAGCAAAAG 279
QY 85  ATTAACCTCGGATTTGGTCTTATAAGATGAGACGGGCAAAACCCCGTACTGACGAGC 144
DB 280 ATGAATCTGGGAGTTGGTGCCTACCGGGATGATAACGGAAAGCCCTTACGTGCTGCTAGC 339
QY 145 GTGAAAAAGGCTGAACAGTATCTGCTCGAAAATGAACCAACCAAAATTTACCTCGGCATT 204
DB 340 GTCCGCAAGGACAGGCCAGATTCGCCGCAAAATTTGGACAGGAATACCTGCCCATTT 399
QY 205 GACGGCATCCCTGAATTTGGTGGTGCATCTCAGGAACCTGCTGTTTGGTAAAGGTAGCGCC 264
DB 400 GGGGACTCGGTGAATTTTCAAGGACATCTCAGAACTAGCCCTGGGTGAGAACAGCGAA 459
QY 265 CTGATCAATGACAAAGCTGCTCGACGCGACAGACTCGGGGGGACACTGGGCGCACTAGC 324
DB 460 GTCTTGAAGAGTGGCGGTTTGTCTACTGTGCAGACCAATTTCTGGAACCTGGAGCCCTTAA 519
QY 325 GTGGCTGCGGATTTCTCGCAAAATACACAGCTTAAGCG---TGTGGGTGAGCAAC 381
DB 520 ATCGGAGCAGTTTCTGCAAGATTTTTTAAGTTCAGCCAGATGCTTTCTGCGCCAA 579
QY 382 CCAAGCTCGCGCAACCATAGAGCGTCTTTAACTCTGAGGTCTGGAAGTTCTGGAATAC 441
DB 580 CCATCTTGGGAAACACACACCCATCTTCAGGGATGCTGGCATGCAGCTACAAAGTTAT 639
QY 442 GCTTATTATGATCGGAAATACACTCTTGACTTCGATGCACTGATTAACAGCCTGAAT 501
DB 640 CGGTATTATGACCCCAAGACTTGGCGTTTTCACCTTCAGAGTGTGTGGAGATATTCA 699
QY 502 GAAGCTCAGGCTGGCGAGCTAGTGTGTTTCAATGCTGCTGCATAACCCCAACCGGTATC 561

```

```

Db 700 AAAATACCAGAGCAGAGTGTCTTCTTTCATGCTGCGCCCAACATCCACGGGAGT 759
QY 562 GACCTACGCTGGAACAATGGCAACACATCTCGTTCGTTGAGAAGGCTGTGTTA 621
Db 760 GACCCGCTCGGAAACAGTGAAGGAATAGCAACAGTGTGTGAAGAAAGGAATCTCTTT 819
QY 622 CCCTGTTTTCATTCGCTTACCAAGGTTTTCGCCGTGG---TCTGGGAAGAATGCTGAA 678
Db 820 GCGTCTTTGACATGGCTTACCAAGGCTTTGCGAGTGGTGATGGTGATAGGATGCGCTGG 879
QY 679 GGAATGGCGCTTTGGGGCTATGCATAAAGAGCTGATTTGTTGCCAGTTCTTACTCTAAA 738
Db 880 GCTGTGCGCCACTTCATCGAACAGGGCATTAATGTTTGTCTCTGCCAATCATATGCCAAG 939
QY 739 AACTTTGGCCTGTACACAGAGCTGTGTGGCGCTTGTACTCTGTTGTCGCGACAGTGA 798
Db 940 AACTGGGCTTATATGGTGAGCGTGTAGGAGCCTTCACTATGTTCTGCAAGATGCGGAT 999
QY 799 ACGTGTGATCGCGCATTCAGCCAAATGAAGCGGCGATTCGCGTAACTACTCTAACCCA 858
Db 1000 GAAGCCAAAGGGTAGAGTCACAGTTGAAGATCTTGATCGTCCCATGTATTCACACCT 1059
QY 859 CAGACACACGGCGCTTCTGTTTGGCCACCATCTCTGAGCAACGATGCTTACGTGCGATT 918
Db 1060 CCCCTCAATGGGCGCGGATTTGCTGTGCTTCTGAACACCCCGAGATTTGCG-AAAAA 1118
QY 919 TGGGAACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGTATGCGTCACTGTTGTCGTC 978
Db 1119 TGGCTGCAAGAGTGAAGGCAATGGCGGACCGGATCAATGGCATGCGGACTCAACTGCTC 1178
QY 979 AATACGCTGCAGGAAAAAGCGCAACCGCGACTTCAGCTTTATCATCAACAGAACGGC 1038
Db 1179 TCAACCTCAAGAGGAGGTTCCACCCACAACACTGGCAACACATCCAGCAAAATTTGGC 1238
QY 1039 ATGTTCTCTTCAGTGGCTGACAAAGAAAGAGTGTGCTGCTGCGCGAGAGTTTGGC 1098
Db 1239 ATGTTCTGTTTACAGGGCTTAAAGCCCTGAACAGGTGGAGCGGCTGATCAAGGAGTTTCC 1298
QY 1099 GTATATGCGGTTGCTTCTGCTGCGTAAATGTGCGCGGATGACACAGATTAACATGCT 1158
Db 1299 ATCTACATGACAAAGGATGGCGCATCTCTGTGGAGGGTCACTTCCAGCAACGTTGGC 1358
QY 1159 CCGCTGTGCGAAGCGATT 1176
Db 1359 TACCTTCCCATGCCATT 1376

RESULT 13
CNS0A2EC 1599 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTGH60ZA09 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX828177
VERSION 1
KEYWORDS HTS: GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 1599)
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche.M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1599)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :

```


BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
<http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis>.

FEATURES

source

1. 1599
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTPGH60A09"
/tissue_type="Hormone Treated Callus"
/ecotype="Col-0"
/plasmid="pCMVSPORT 6"
complement(1..1599)
/gene="At4g31990"

gene

ORIGIN

Query Match 17.2%; Score 205.4; DB 4; Length 1599;
Best Local Similarity 49.5%; Pred. No. 2.1e-52;
Matches 586; Conservative 0; Mismatches 591; Indels 6; Gaps 2;

QY 4 TTGAGACATTACCGCGCTCTCGCGACCGATTCTGGCGCTCGCGCATCTGTTTGGT 63
DB 218 TTCAGGGTAACTATCGCTCCACGACCCCTATCTTGAGTCAGTTAAGCATCAAA 277

QY 64 GCCGATGAACCTCCCGCAAAATTAACCTCGGGATTGGTGTCTATAAAGATGAGACGGGC 123
DB 278 GCTGACACTAAGCGGATGAACCTCAATCTTGGTGTGGTCTTATCGTACTGAGGAATC 337

QY 124 AAAACCCGGTACTGACGAGGTGAAAAGGTGAAAGGTGAAAGTATCTGCTCGAAATGAACC 183
DB 338 CAGCCTTATGTGCTTAATGTTGTTAAAGGCGGAGAAATTTGATGTTGGAGAGGAGAT 397

QY 184 ACCAAAAATTACCTCGGCATTGACGGCATCCTGAAATTTGTCGTCGCTCAGCACTG 243
DB 398 AACAAAGATATCTTCCAAATGAGGGGTGGCAGCATTCACAGGCTACTGCTGAGTTG 457

QY 244 CTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTGCTCGCAGCGCACAGACTCCG 303
DB 458 CTATTTGGAGCTGGTCTATCTCTGTTATTAAAGGAACAAGAGTAGCAACAATTCAGGGTCT 517

QY 304 GGGGCACTGGCGCACTAGCGTGGCTCCGATTTCTGGCAAAATAATACGAGCGTTAAG 363
DB 518 TCGGGAACAGGTTCACTCGGATAGCAGCGCTCTTATAGACGCTTATTCCTCGAGCA 577

QY 364 CGTGTGGGTGAGCAACCCAGCTGGCGACCATAGAGCGTCTTAACTCTGCGAGT 423
DB 578 AAAGTTGTGATCTCATCAACCACTGGGGTAATCACAAGATATCTTCAATGATGCCAAA 637

QY 424 CTGGAAGTTTCGTGAATACGCTTATATGATCGGAAATACACCTTGTGACTTCGATGCA 483
DB 638 GTTCCGTGTCGGAATACCGCTACTATGATCCAAACAAATTTGGTATTTGAGGGA 697

QY 484 CTGATTAACAGCTGAATGAAGCTCAGGCTGGGACGTPAGTGTGTTTCCATGGCTGTGC 543
DB 698 ATGATAGCAGATATAAAGGAAGCTCCAGAAGGATCCTTCATCTTGTCTCACGGATGTCT 757

QY 544 CATTAACCAACCGGTATCGACCTTACGCTGGAACAATGGCAACACTGGCAACTCTCC 603
DB 758 CACAAACCACTGAAATGACCCCAACACCAAGACAGTGGGTAAATAATGCTGATGTATT 817

QY 604 GTTGAGAAAGGCTGTTACCGCTGTTTACCTGTTGACTTTCGCTTACCAGGGTTTTGCCCGTG--GT 660

DB 818 CAGGAAAGAACCATATCCCATTTTTCGATGTCGATACCAAGGCTTTGCTAGTGAAGC 877
QY 661 CTGGAAGAAGATGCTGAAGGACTCGCGCTTTTCGGGCTATGCATAAAGAGCTGATTGT 720
DB 878 CTTGATGAAGATGAGCAGCATCTGTGAGATATTTCGTGAGCGGGGAATGGAGTTTGT 937
QY 721 GCCAGTTCTACTCTAAAAACCTTTGGCCCTGTACACGAGCGTGTGGCGCTTTGTACTCTG 780
DB 938 GCTCAGTCATATAGTAAAAATTTAGTTTGTATGCAGAAAGAAATTTGGGCAATCAATGTC 997
QY 781 GTTGTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGACGCGGATTCGC 840
DB 998 GTGTGCTCCTCAGCTGATGCTGTACAGGGGTCAAGGCCAGTTTGAAAAGGATTTGCTCGG 1057
QY 841 GCTAACTACTCTAAACCCACGACACGCGCTTCTGTGTTGCCACCATCTCTGAGCAAC 900
DB 1058 CTTATGTAATCCAGTTCATGGGGGAGATCGTGGCCAAATGTAGTGGGTGAT 1117
QY 901 GATCGCTTACGTCGATTTGGGAAACAAGAGCTGATGATGCGCCAGCGTATTTCAGCGT 960
DB 1118 GTAATATGTTTCAGTGAATGGAAGCAGAGATGGAATGATGGCAGGAAGATTAAGACG 1177
QY 961 ATGCGTCAGTTGTTGTCATACGCTGCAGGAAAAAGCGCAACCG--CGACTTCAGC 1017
DB 1178 GTTAGACAAGAGCTGTATGATAGCTCGTTTCAAAGACAAGAGCGGAGGACTGGTCC 1237
QY 1018 TTTATCATCAACGAGCGGCTGTTCTCTTCAGTGGCTGACACAAAAGACAAAGTGTG 1077
DB 1238 TTCATCTGAAGCAAAATGGGCATGTTCTCTTTCACCGGCCCTAAACAAAGCTCAGAGC 1297
QY 1078 CGTCTGCGCAAGAGTTTGGCGTATATGCGGTGCTTCTGCTGCGGTAATGTTGCCCGG 1137
DB 1298 AACATGAGCAAAATGGCATGTGTATGATTAAGACGGGAGATATCATTTGSCCGGA 1357
QY 1138 ATGACACAGATAACATGGCTCCGCTGTGCGAAGCGATTGTGG 1180
DB 1358 TTATCTCTGGCAAAATCGAGTATCTTGTCTGATCGATCATCG 1400

RESULT 14

CK259742

LOCUS

DEFINITION EST705820 potato abiotic stress cDNA library Solanum tuberosum cDNA
clone POAB141 5' end, mRNA sequence.

CK259742

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. 977
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POAB141"
/tissue_type="abiotic stress treated leaf and root tissue"

CK259742 977 bp mRNA linear EST 03-AUG-2004
EST705820 potato abiotic stress cDNA library Solanum tuberosum cDNA
clone POAB141 5' end, mRNA sequence.

CK259742
EST705820
GI:39816720

Solanum tuberosum (potato)
Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum.

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
Generation of ESTs from abiotic stressed potato tissue

Unpublished (2003)
Other_ESTs: EST705819 EST705821

Contact: Robin Buell
The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics
Institute via <http://genome.arizona.edu/orders/>.

Seq primer: CAG GAA ACA GCT ATG ACC.
Location/Qualifiers

```
/lab host="DHL0B-TonA"
/clone_lib="potato abiotic stress cDNA library"
/notes=vector: pCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."
```

ORIGIN

Query Match	17.0%;	Score 203;	DB 7;	Length 977;
Best Local Similarity	51.3%;	Pred. No. 1e-51;		
Matches 498;	Conservative	0;	Mismatches 470;	Indels 3; Gaps 1;
Qy	208	GGCATCCCTGAATTTGGTGCCTGCACTCAGGAACTGCTGTTTGGTAAGGTAGGCGCCCTG	267	
Db	2	GACTTGCTGGATTTTAAACAAATTCAGTGCACCAAGCTTATATTGGTGTGACAGCGCCCTGCC	61	
Qy	268	ATCAATGACAACTGCTCGCACGGCAGACCTCGGGGGGCACTGGCGCACTACGGGTG	327	
Db	62	ATTAAGAGAAATAGGGTGACCACTGTTCAAGTCTGTTGCGGCACTGGTCTTTGAGGGTT	121	
Qy	328	GCTGCGGATTTCTGGCAAAAATACCAAGCGCTTAAGCGTGTGTTGGGTGAGCAACCCCAAGC	387	
Db	122	GGTGTGATTTCTGGCCAGGCAITATCATGAACATATATATATACCAAGCAACA	181	
Qy	388	TGGCGGAACCAATAGAGCGCTTTAACTCTCGAGGCTCTGGAAGTTCGTAATACGCTTAT	447	
Db	182	TGGGGAACCACTCTTAAGATTTTCACTTTAGCTGGGCTTCAGTAAATAATATTATCGTTAC	241	
Qy	448	TATGATCGGAATATCACACTCTTGACTTCGATCGACTGATTACAGCGCTCAATGAAGCT	507	
Db	242	TATGACCCAGCAACACGAGGCGCTAGATTTTCAAGGATTTATGGATGATCTTACGGCTGCA	301	
Qy	508	CAGGCTGGGAGCTAGTCTGTTCATGGCTGCTGCCATAACCCACCGGTATCGACCCCT	567	
Db	302	CCCTCTGTGCAATAGTTCTTCTCATCTGTTGCTCTATACCCCACTGGCGTTGATCCA	361	
Qy	568	ACGCTGGAACCAATGCAACACACTGGCACAACTCTCCGTTGAGAAGGCTGTTACCGCTG	627	
Db	362	ACAAATTGACCAAGTGGGAAAAAATCAGGAAGTTGATGAGGTCACAGGGGTTGTACCTTTC	421	
Qy	628	TTTGACTTCGCTTACCAAGGTTTGGCCGCTGG---TCTGGAAGAGATGCTGAAGGACTG	684	
Db	422	TTTGACAGTGTATTATCAGGGTTTTTGGCCAGTGGAAACCTAGATGCAATCTGTT	481	
Qy	685	CGCGTTTTCGGGCTTATGCATAAAGAGCTGATTTGTGCGAGTCTCTACTCTTAAAACTTT	744	
Db	482	CGCATGTTTGGCTGTAGTGGTGAATGTCTTGACAGCTCAGAGTTATGCCAAAAACATG	541	
Qy	745	GGCCTGTAAACAGAGCGTGTGGCGCTTGTACTCTGTTGTCGCGACAGTGAACCGCTT	804	
Db	542	GGACTGTATGAGAGCGTGTGGTCCCTTACGATTTGTGCAAAAAATAGCAATGTTTGA	601	
Qy	805	GATCGCGATTCAGGCCAATGAAGCGCGGATTCGCGCTACTCTTAACCCACCAAGCA	864	

Db	602	AGCAAAGTCGAAAGCCAGCTAAAGATAGTTTATAAGGCCAATGTACTCAAATCCACCGATT	661	
Qy	865	CACGGCGCTTCTGTTGTTGCCACCATCTCTGAGCAACGATCGGTACGTGGAITTTGGAA	924	
Db	662	CATGGTGCCTCTATTGTTGCCACTATCTCAAGGACAGACAAAATGTATGACGAATGGACA	721	
Qy	925	CAAGAGCTGACTCATATGCGCCAGCGTATTATCAGCGTATGCGTTCAGTTGTTTCGTCATACG	984	
Db	722	GTTGAGCTGAAGCAATGCGCGACCGAATTTATTAGCATGCGCCAGCAACTCTATGATGCC	781	
Qy	985	CTCAGGAAAAAGCCGCAAAACCCGCACTTCAGCTTTATCATCAAAAGAGCGGATGTTTC	1044	
Db	782	ATCCAAGCTCAAGGTACTCCAGGTGATTGGAGTCATATCATCAAGCAAAATTTGGATGTTT	841	
Qy	1045	TCCTTGTGCGCTTGACAAAAGAACAAAGTCTGCTGCTGCGGAGAGTTTGGCGTATAT	1104	
Db	842	ACTTTCCACAGGATTTGAATGTTTGACCAAGTCTCGTTCATGACTAGAGAGTATCACATTTAC	901	
Qy	1105	GCGGTTGCTTCTGCTCGCGTAAATGTGGCGGATGACACAGATACATATGCGTCCGCTG	1164	
Db	902	ATGACATCTGATGGGAGAAATTAGCATGGCTGGTCTTAGTTCTCGGACAGTTTCTCATCTT	961	
Qy	1165	TGCGAAGCGAT 1175		
Db	962	GCTGATGCCAT 972		

RESULT 15	CNS092WQ	1486 bp	mRNA	linear	HTC 06-FEB-2004
LOCUS	Arabidopsis thaliana Full-length cDNA Complete sequence from clone				
DEFINITION	GS17FB192F03 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).				
ACCESSION	BX829537				
VERSION	1	GI:42456349			
KEYWORDS	HTC; GSLT cDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.				
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1486)				
AUTHORS	Genoscope.				
JOURNAL	Direct Submission				
TITLE	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr				
COMMENT	- Web : www.genoscope.cns.fr) The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full-length http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.				
FEATURES	Location/Qualifiers				
source	1. .1486				
	/organism="Arabidopsis thaliana"				
	/mol_type="mRNA"				
	/db_xref="taxon:3702"				

```

/clonename="GSJTFB19ZF03"
/tissue_type="Flowers and buds"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1. .1486
/gene="At5g19550"

```

gene

ORIGIN

Query Match	16.8%	Score 200.6;	DB 4;	Length 1486;
Best Local Similarity	50.9%;	Pred. No. 6.6e-51;		
Matches 605;	Conservative 0;	Mismatches 574;	Indels 10;	Gaps 5;
QY	10	AACATTACCGCGCTCTCGCGACCCCGATTCTGGGCCCTGGCCGATCTGTTTCGTGCGCAT	69	
DB	108	AACGTTGCTGTGCTCCGGAAGATCCTATTCTCGGTGTAACTGTTGCTTTACAACAATGAT	167	
QY	70	GAACGTCCTCCGCAAAATTAACCTTCGGATTGGTGTCTATAAAGATGAGACGGGCAAAACC	129	
DB	168	CCTAGTCTCTGTTAAGATCAATTTGGGTGTCGTCCTACCGAACTGAGGAAGGAAGCCT	227	
QY	130	CCGGTACTGACAGCGTGAAAGAGCTGAACAGTATCTGTCGAAATGAAACCAAC	186	
DB	228	CTTTGTTCTTGATGTGGTGCCTGTAAGGCAGAGCAACAGCTAGTGAACGATCCGCTCTCGGGTC	287	
QY	187	AAAAATTACCTCGGCATTTGACGGCATCCCTGAAATTTGGTTCGCTGCACCTCAGGAACCTGCTG	246	
DB	288	AAGGAATACATTCCTATGTTGGTATTTCCGATTTTAAAGTTAAGGCCCAAGCTCATC	347	
QY	247	TTTGGTAAAGTAGCGCCCTGATCAATGACAAACGTCCTCGCAGGCACACAGACTCCGGGG	306	
DB	348	TTAGGTGCTGACAGTCTCGGATTTACAGAAAGTAGAGTTACTACAGTCCAGTCTGTCT	407	
QY	307	GGCACTGGCGCACTACGCGTGGCTGCCGATTTCTTGGCCAAATAATACAGGGTTAAGCGT	366	
DB	408	GGTACTGGTTGCTTGAAGATTGGGGCTGAGTTCCTCAAAACACACTACCACCAAAAGTGC	467	
QY	367	GTGTGGGTGAGCAACCAAGCTGGCCGAACCAT-AAGAGCGCTCTTTAACTCTGCAGGTCT	425	
DB	468	ATTTACATTTCCAAACCAACTTGGGGAAACATCCCAAGTTTCAACTTGGCAGGCTT	527	
QY	426	GSAAAGTTCGTGAATACGCTTATTATGATGCGGAAATCACACTCTTGACTTCGATGCACT	485	
DB	528	GTCTGTGGAGTATTTCCGCTACTATGATCCTGCACCCCGTGTCTTGACTTCAAGGCTT	587	
QY	486	GATTAACAGCCTGAAATGAAGCTCAGGCTGGCGAGTAGTGTGTTCCATGGCTGCTGCCA	545	
DB	588	GCTCGAGGATCTTGGCGCTGCACCATCTGAGAGCTATTGTCTTACTTCATGCATGTTCCACA	647	
QY	546	TAAACCCAAACGGGTATCGACCTCCTACGCTGGAAACAATGCAAAACACTGCGCAACAACCTCCCGT	605	
DB	648	CAATCCCACTGGAGTTGACCCCACTCTGAACAGTGGGAACAGATTCGAAGCTGATGAG	707	
QY	606	TGAAAGGCTGTTTACCGCTGTTTGACTTTCGCTTACAGGGTTTTCGCCGTGT---	662	
DB	708	ATCTAAAGCTTATTACCCCTTTTGTATAGTGCATATCAGGGTTTTCGTAGTGGTAGCCT	767	
QY	663	GGAGAGATGCTGAMGGACTGCGCGCTTTCGCGGCTATGCATAAAGAGCTGATTGTTGC	722	
DB	768	TAAACACAGATGCAACAATCTGTCGGTCAATTTGTTGCTGATGGCGGTGAATGCTTGTATAGC	827	
QY	723	CAGTTCTACTCTAAAAAATTTGGCCCTGTACAACGAGCGTGTGGCGCTTGTACTCTGGT	782	
DB	828	TCAAGTTACGCCCAAAATATGGACTTTATGGGAGCGTGTGGTGGCTTAGCATTTGT	887	
QY	783	TGCTGCGGACAGTGAAACCGTTGATCCGCATTCAGCCAAATGAAAGCGCGGATTTGCGGC	842	
DB	888	CTCAAGTCAAGCAGATGTGGCTAGTAGGTTGAGAGCCAGGTGAAGCTTGTGTGCGGCC	947	
QY	843	TAACTACTCTAAACCCACAGCA-CAGGGCCCTTCTGTTGTTGCCACCATCCTGAGCAACG	901	
DB	948	CATGTATTGAGCCCCCCCCTATTTCATGGAGCATCAATTGTTGGCACCATTTCTAAAAAGCA	1007	
QY	902	ATCGGTTACGTGCGAATTTGGGAAACAAGAGCTGACTGATATGCGCCACGCTATTCA--GCG	959	

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2006, 13:59:01 ; Search time 6319 Seconds
(without alignments)
10713.808 Million cell updates/sec

Title: US-10-673-786A-1

Perfect score: 1191
Sequence: 1 atgttgagaacattaccgc.....cgattggcagtgctgtaa 1191

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_in: *
3: gb_env: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pr: *
9: gb_ro: *
10: gb_sts: *
11: gb_sy: *
12: gb_un: *
13: gb_vi: *
14: gb_htg: *
15: gb_pl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1191	100.0	1293	6	I08485 Sequence 12
2	1191	100.0	1331	6	X05904 E. coli asp
3	1191	100.0	1331	6	AR095441 Sequence
4	1191	100.0	1331	6	E16758 gDNA encodi
5	1191	100.0	1415	1	X03629 E. coli asp
6	1191	100.0	1468	6	I08484 Sequence 9
7	1191	100.0	3659	6	I08487 Sequence 15
8	1191	100.0	11470	1	D90731 Escherichia
9	1191	100.0	14759	6	AX370270 Sequence
10	1191	100.0	20604	1	D90730 Escherichia
11	1191	100.0	110000	1	Continuation (10 o
12	1170.2	98.3	110000	1	Continuation (12 o
13	1170.2	98.3	110000	1	Continuation (13 o
14	1170.2	98.3	110000	1	Continuation (12 o
15	1168.6	98.1	110000	1	Continuation (10 o
16	1168.6	98.1	292504	1	AE016981 Shigella
17	1159	97.3	301276	1	AE016758 Escherich
18	864.6	72.6	110000	1	Continuation (19 o

C 19	861.4	72.3	62027	1	AE008743	AE008743 Salmonell
C 20	861.4	72.3	294050	1	AL627268	AL627268 Salmonell
C 21	861.4	72.3	301983	1	AE016840	AE016840 Salmonell
C 22	855	71.8	110000	1	AE017220_10	Continuation (11 o
C 23	839	70.4	1209	6	AR388173	AR388173 Sequence
C 24	743	62.4	110000	1	EX936398_17	Continuation (18 o
C 25	741.4	62.3	14855	1	AE013879	AE013879 Yersinia
C 26	741.4	62.3	210050	1	AJ414148	AJ414148 Yersinia
C 27	741.4	62.3	291847	1	AE017131	AE017131 Yersinia
C 28	728.6	61.2	110000	1	EX950851_28	Continuation (29 o
C 29	691.8	58.1	249878	6	AX770903	AX770903 Sequence
C 30	691.8	58.1	349071	1	EX571864	EX571864 Photornab
C 31	589	49.5	589	1	AY698247	AY698247 Escherich
C 32	589	49.5	589	1	AY698283	AY698283 Escherich
C 33	585.2	49.1	590	1	AY698263	AY698263 Shigella
C 34	584	49.0	584	1	AY698244	AY698244 Escherich
C 35	584	49.0	584	1	AY698246	AY698246 Escherich
C 36	583	49.0	583	1	AY698284	AY698284 Escherich
C 37	582.2	48.9	587	1	AY698252	AY698252 Shigella
C 38	581.6	48.8	588	1	AY698255	AY698255 Shigella
C 39	581.4	48.8	583	1	AY698248	AY698248 Escherich
C 40	581.2	48.8	586	1	AY698280	AY698280 Escherich
C 41	580.6	48.7	587	1	AY698287	AY698287 Shigella
C 42	580.6	48.7	587	1	AY698288	AY698288 Shigella
C 43	580.4	48.7	590	1	AY698269	AY698269 Shigella
C 44	580.2	48.7	585	1	AY698261	AY698261 Shigella
C 45	580.2	48.7	585	1	AY698266	AY698266 Shigella

ALIGNMENTS

RESULT 1
I08485
LOCUS I08485 1293 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 12 from Patent WO 8700202.
ACCESSION I08485
VERSION I08485.1 GI:588805
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1293)
AUTHORS Edwards, M.R., Taylor, P.P., Hunter, M.G. and Fotheringham, I.G.
TITLE COMPOSITE PLASMIDS FOR AMINO ACID SYNTHESIS
JOURNAL Patent: WO 8700202-A 12 15-JAN-1987;
FEATURES
source
1. .1293
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match	100.0%;	Score 1191;	DB 6;	Length 1293;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1191;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	ATGTTTGAACATTACCGCGCTCTGCGACCGCATTCCTGGCCCTGGCGGATCTGTTT	60	
Db	12	ATGTTTGAACATTACCGCGCTCTGCGACCGCATTCCTGGCCCTGGCGGATCTGTTT	71	
QY	61	CGTGCCGATGAACGCTCCCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG	120	
Db	72	CGTGCCGATGAACGCTCCCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG	131	
QY	121	GGCAAAACCCGGTACTGACCGCTGAAAGGCTGAAACAGTATCTGTCGAAAATGAA	180	
Db	132	GGCAAAACCCGGTACTGACCGCTGAAAGGCTGAAACAGTATCTGTCGAAAATGAA	191	
QY	181	ACCACAAAATTACCTCGGATTTGGTGTCTATAAAGATGAGACG	240	
Db	192	ACCACAAAATTACCTCGGATTTGGTGTCTATAAAGATGAGACG	251	
QY	241	CTGCTGTTTGGTAAAGGTAGCGCCTGATCAATGACAAACGCTGCTGCGACGACAGCT	300	

```
Db      252  CTGCTGTTTGGTAAAGTAGCGCCCTGATCATATGACAAAGCTGCTCGACGGACAGACT 311
Qy      301  CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGCAAAAAATACCAACGTT 360
Db      312  CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGCAAAAAATACCAACGTT 371
Qy      361  AAGCGTGTGGGTGAGCAACCCCAAGCTGGCCGCAACCATAGAGCGTCTTAACTCTGCA 420
Db      372  AAGCGTGTGGGTGAGCAACCCCAAGCTGGCCGCAACCATAGAGCGTCTTAACTCTGCA 431
Qy      421  GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAATACACACTCTTGACTTCGAT 480
Db      432  GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAATACACACTCTTGACTTCGAT 491
Qy      481  GCACTGATTAAACAGCCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGCTGC 540
Db      492  GCACTGATTAAACAGCCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGCTGC 551
Qy      541  TGCATTAACCCAAACCGGTATCGACCCCTACGCTGGAACAATGGCAACACTGGCACAAC 600
Db      552  TGCATTAACCCAAACCGGTATCGACCCCTACGCTGGAACAATGGCAACACTGGCACAAC 611
Qy      601  TCCGTTGAGAAAGCGTGGTTACCGCTGTTTGAAGCTTACGCTTACCAAGGTTTTCGCCGTGG 660
Db      612  TCCGTTGAGAAAGCGTGGTTACCGCTGTTTGAAGCTTACGCTTACCAAGGTTTTCGCCGTGG 671
Qy      661  CTGGAAGAAGATGCTGAAGGACTCGCGCTTTTCGCGCTATGCTATGAAGAGCTGATGTT 720
Db      672  CTGGAAGAAGATGCTGAAGGACTCGCGCTTTTCGCGCTATGCTATGAAGAGCTGATGTT 731
Qy      721  GCCAGTTCCTACTCTATAAAACCTTTGGCGCTGTACAAAGAGCGTGTGGCGTTTGTACTCTG 780
Db      732  GCCAGTTCCTACTCTATAAAACCTTTGGCGCTGTACAAAGAGCGTGTGGCGTTTGTACTCTG 791
Qy      781  GTTGCTGCCACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTTCG 840
Db      792  GTTGCTGCCACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTTCG 851
Qy      841  GCTAACTACTCTAACCCACAGCACAGCGGCTTCTGTTGTCGACCATCTTGAGCAAC 900
Db      852  GCTAACTACTCTAACCCACAGCACAGCGGCTTCTGTTGTCGACCATCTTGAGCAAC 911
Qy      901  GATGCGTTACGTGCGATTGGGAACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960
Db      912  GATGCGTTACGTGCGATTGGGAACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 971
Qy      961  ATGCGTCAGTTGTTGTCGTCATACGCTGCAGGAAAAAGGCGCAAAACCGCACTTCAGCTTT 1020
Db      972  ATGCGTCAGTTGTTGTCGTCATACGCTGCAGGAAAAAGGCGCAAAACCGCACTTCAGCTTT 1031
Qy      1021  ATCATCAACAGAACGCGCATGTTCTCTTCAGTGGCCTGACAAAGAACAAAGTGTGCGT 1080
Db      1032  ATCATCAACAGAACGCGCATGTTCTCTTCAGTGGCCTGACAAAGAACAAAGTGTGCGT 1091
Qy      1081  CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGCGTGGTAAATGTCGCGGGATG 1140
Db      1092  CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGCGTGGTAAATGTCGCGGGATG 1151
Qy      1141  ACACCATACATGCTCCGCTGTGCGAAGCGATTGTCGAGTGCCTGTA 1191
Db      1152  ACACCATACATGCTCCGCTGTGCGAAGCGATTGTCGAGTGCCTGTA 1202
```

```
RESULT 2
ECASPCG  ECASPCG 1331 bp DNA linear BCT 12-SEP-1993
LOCUS    E. coli aspC gene for aspartate aminotransferase.
DEFINITION X05904
ACCESSION X05904
VERSION X05904.1 GI:41012
KEYWORDS  aspartate aminotransferase; aspC gene.
SOURCE    Escherichia coli
ORGANISM  Escherichia coli
```

```
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 1331)
Kuramitsu.S., Okuno.S., Ogawa,T., Ogawa,H. and Kagamiyama,H.
Aspartate aminotransferase of Escherichia coli: nucleotide sequence
of the aspC gene
J. Biochem. 97 (4), 1259-1262 (1985)
PUBMED 3897210
2 (bases 1 to 1331)
Kagamiyama,H.
Direct Submission
Submitted (07-MAR-1988)
Data kindly reviewed (07-MAR-1988) by Kagamiyama H.
Location/Qualifiers
source
1..1331
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
10..1200
/note="unnamed protein product; aspartate aminotransferase
(AA 1 - 396)"
/codon_start=1
/transl_table=11
/protein_id="CAA29333.1"
/db_xref="GI:41013"
/db_xref="GOA:P00509"
/db_xref="UniProt/Swiss-Prot:P00509"
/translation="MFENITAAPADPILGLADLFRADPRPGKINLIGIYVKDTGKTP
VITSVKAAQDYLLNETTKYLGIDGIPFGFCTQELLFGKSGALINDKRARTHTDPTG
CTGALRVADFLAKNTSVKRVVSNPNSPHKSVFNSAGLEVREYAYDAENHTLDFD
ALINSLEAQAQGVVLFHGCCHNPTGIDTLPQWOTLAQSLVEKWLPLDFPAYQGFPA
RGLEDAELGRAPAAHMKELIVASSYKSNFGLYNERVGACTLVAADSETVDRAFSQMK
AAIRANYSNPPAHGSVVATILSNDAIRAIWEQELTDMRQIORMRQLFVNTLQEKGA
NDFSFIIKQNGWFSFGTLKEQVLNREEFGVYAVASRVNVAGTPTPDNMAPLCEAI
VAVL"
old_sequence 90
/note="c was a in [1]"
/citation=[1]
old_sequence 502..503
/note="ag was tc in [1]"
/citation=[1]
old_sequence 612
/note="c was g in [1]"
/citation=[1]
old_sequence 657
/note="t was g in [1]"
/citation=[1]
old_sequence 693
/note="g was c in [1]"
/citation=[1]
old_sequence 696
/note="c was t in [1]"
/citation=[1]
old_sequence 699
/note="t was g in [1]"
/citation=[1]
ORIGIN
Query Match 100.0%; Score 1191; DB 1; Length 1331;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTTTGAGAA CATTACCGCGCTCTCTGCGACCCGATTCGGGCTTGGCCGATCTGTTT 60
Db 10 ATGTTTGAGAA CATTACCGCGCTCTCTGCGACCCGATTCGGGCTTGGCCGATCTGTTT 69
Qy 61 CGTGCGGATGAACGTCCTCCGCAAAATTAACCTCGGATTCGTTCTATAAAGATGACACG 120
Db 70 CGTGCGGATGAACGTCCTCCGCAAAATTAACCTCGGATTCGTTCTATAAAGATGACACG 129
Qy 121 GCGAAAAACCCCGTACTGACACCGCTGAAAAAGCGTGAAACAGTATCTGCTCGAAATGAA 180
Db 130 GCGAAAAACCCCGTACTGACACCGCTGAAAAAGCGTGAAACAGTATCTGCTCGAAATGAA 189
```

181 ACCACCAAAATTAACCTCGCATTTGACGGCATCCCTGAATTTGGTCTGCTCACTCAGGAA 240
 190 ACCACCAAAATTAACCTCGCATTTGACGGCATCCCTGAATTTGGTCTGCTCACTCAGGAA 249
 241 CTGCTGTTGGTAAAGTAGGGCTGATCAATGACAAAGCTGCTGCGACAGACT 300
 250 CTGCTGTTGGTAAAGTAGGGCTGATCAATGACAAAGCTGCTGCGACAGACT 309
 301 CCGGGGGCACTGCGCACTACCGTGGCTGCGATTTCTGCGCAAAATATACAGGTT 360
 310 CCGGGGGCACTGCGCACTACCGTGGCTGCGATTTCTGCGCAAAATATACAGGTT 369
 361 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 420
 370 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 429
 421 GGTCTGAAATTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGATTCGAT 480
 430 GGTCTGAAATTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGATTCGAT 489
 481 GCACGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTTCCATGGCTGC 540
 490 GCACGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTTCCATGGCTGC 549
 541 TGCCATAAACCAACCGGTATCGACCTTACGCTCGAACCAATGGCAACACCTGGCACAACTC 600
 550 TGCCATAAACCAACCGGTATCGACCTTACGCTCGAACCAATGGCAACACCTGGCACAACTC 609
 601 TCGGTTGAGAAAGCGTGTACCGCTGTTTGAATTCGCTTACAGGGTTTGGCCGCTGGT 660
 610 TCGGTTGAGAAAGCGTGTACCGCTGTTTGAATTCGCTTACAGGGTTTGGCCGCTGGT 669
 661 CTGGAAGAAGATGCTGAAGACTGCGCGCTTTGCGGCTATGCTAAGAGCTGATTTGTT 720
 670 CTGGAAGAAGATGCTGAAGACTGCGCGCTTTGCGGCTATGCTAAGAGCTGATTTGTT 729
 721 GCCAGTTCCTACTCTAAATACTTTGGCTCTGTAACAAGAGCTGTTGGCGCTTGTACTCTG 780
 730 GCCAGTTCCTACTCTAAATACTTTGGCTCTGTAACAAGAGCTGTTGGCGCTTGTACTCTG 789
 781 GTTGCTGCCAGACAGTGAACCGTTGATTCGCGCAATTCAGCCAAATGAAGCGGCGATTCG 840
 790 GTTGCTGCCAGACAGTGAACCGTTGATTCGCGCAATTCAGCCAAATGAAGCGGCGATTCG 849
 841 GCTAACTACTCTAAACCAACAGCAGCGGCTTCTGTTGCTGCCACCATCTCTGAGCAAC 900
 850 GCTAACTACTCTAAACCAACAGCAGCGGCTTCTGTTGCTGCCACCATCTCTGAGCAAC 909
 901 GATGCGTTACGTGCGATTTGGGAACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960
 910 GATGCGTTACGTGCGATTTGGGAACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 969
 961 ATGCGTCAGTGTGTCGTAATAGCTGAGGAAAGCGCAACCGCGACTTCAGCTTT 1020
 970 ATGCGTCAGTGTGTCGTAATAGCTGAGGAAAGCGCAACCGCGACTTCAGCTTT 1029
 1021 ATCATCAACAGAACCGCATGTTCTCTTCAAGTGGCTGACAAAGAACAAAGTGTGCGT 1080
 1030 ATCATCAACAGAACCGCATGTTCTCTTCAAGTGGCTGACAAAGAACAAAGTGTGCGT 1089
 1081 CTGCGGAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTTGGCCGGATG 1140
 1090 CTGCGGAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTTGGCCGGATG 1149
 1141 ACACGAGTAACATGCTCGCTGTGCGAAGCGATTTGTGCGAGTGTCTGTAA 1191
 1150 ACACGAGTAACATGCTCGCTGTGCGAAGCGATTTGTGCGAGTGTCTGTAA 1200

RESULT 3
 AR095441
 LOCUS AR095441 1331 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 23 from patent US 6004773.
 ACCESSION AR095441
 VERSION AR095441.1 GI:10023318
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1331)
 AUTHORS Araki,M., Sugimoto,M., Yoshihara,Y. and Nakamatsu,T.
 TITLE Method for producing L-lysine
 JOURNAL Patent: US 6004773-A 23 21-DEC-1999;
 FEATURES
 source Location/Qualifiers
 1..1331
 /organism="unknown"
 /mol_type="unassigned DNA"

Query Match 100.0%; Score 1191; DB 6; Length 1331;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTTGAGAACTTACCGCGCTCTGCGACCCCGATTCGCGCCTGGCGGATCTGTTT 60
 Db 10 ATGTTTGAGAACTTACCGCGCTCTGCGACCCCGATTCGCGCCTGGCGGATCTGTTT 69
 Qy 61 CGTGCGGATGAACGCTCCCGCAAAATTAACCTCGGATTTGTTCTATAAGATGAGACG 120
 Db 70 CGTGCGGATGAACGCTCCCGCAAAATTAACCTCGGATTTGTTCTATAAGATGAGACG 129
 Qy 121 GGCACCAACCCCGTACTGACAGCGTGAAAGAGCTGAAACAGTATCTGCTCGAAATGAA 180
 Db 130 GGCACCAACCCCGTACTGACAGCGTGAAAGAGCTGAAACAGTATCTGCTCGAAATGAA 189
 Qy 181 ACCACCAAAATTAACCTCGGATTTGAGCGGATCCCTGAATTTGGTCTGCTCAGGAA 240
 Db 190 ACCACCAAAATTAACCTCGGATTTGAGCGGATCCCTGAATTTGGTCTGCTCAGGAA 249
 Qy 241 CTGCTGTTGGTAAAGTAGCGCCTGATCAATGACAAACAGCTGCTGCGGACAGACT 300
 Db 250 CTGCTGTTGGTAAAGTAGCGCCTGATCAATGACAAAGCTGCTGCGGACAGACT 309
 Qy 301 CCGGGGGCACTGCGGCACTACGCTGGCTGCGATTTCTGCGCAAAATATACAGGTT 360
 Db 310 CCGGGGGCACTGCGGCACTACGCTGGCTGCGATTTCTGCGCAAAATATACAGGTT 369
 Qy 361 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 420
 Db 370 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 429
 Qy 421 GGTCTGAAATTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGATTCGAT 480
 Db 430 GGTCTGAAATTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGATTCGAT 489
 Qy 481 GCACGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTTCCATGGCTGC 540
 Db 490 GCACGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTTCCATGGCTGC 549
 Qy 541 TGCCATAAACCAACCGGTATCGACCTTACGCTCGAACCAATGGCAACACCTGGCACAACTC 600
 Db 550 TGCCATAAACCAACCGGTATCGACCTTACGCTCGAACCAATGGCAACACCTGGCACAACTC 609
 Qy 601 TCGGTTGAGAAAGCGTGTACCGCTGTTTGAATTCGCTTACAGGGTTTGGCCGCTGGT 660
 Db 610 TCGGTTGAGAAAGCGTGTACCGCTGTTTGAATTCGCTTACAGGGTTTGGCCGCTGGT 669
 Qy 661 CTGGAAGAAGATGCTGAAGACTGCGCGCTTTGCGGCTATGCTAAGAGCTGATTTGTT 720
 Db 670 CTGGAAGAAGATGCTGAAGACTGCGCGCTTTGCGGCTATGCTAAGAGCTGATTTGTT 729
 Qy 721 GCCAGTTCCTACTCTAAATACTTTGGCTCTGTAACAAGAGCTGTTGGCGCTTGTACTCTG 780
 Db 730 GCCAGTTCCTACTCTAAATACTTTGGCTCTGTAACAAGAGCTGTTGGCGCTTGTACTCTG 789

Db 970 ATCGGTGAGTGTGTCATACGCTGACGAGAAAAAGGCGCAAAACCGGACCTTCAGCTTT 1029
 QY 1021 ATCATCAAAACAGACGCGCATGTTCTCTCTTCAGTGGCTGACAAAAGAAACAAGTGTGCGT 1080
 Db 1030 ATCATCAAAACAGACGCGCATGTTCTCTCTTCAGTGGCTGACAAAAGAAACAAGTGTGCGT 1089
 QY 1081 CTGCGGGAAGTTCGCGGTATATGCGGTGCTCTCTCTTCAGTGGCTGACAAAAGAAACAAGTGTGCGGATG 1140
 Db 1090 CTGCGGGAAGTTCGCGGTATATGCGGTGCTCTCTCTTCAGTGGCTGACAAAAGAAACAAGTGTGCGGATG 1149
 QY 1141 ACACAGATTAACATGCTCGCTGTCGAGCGATTTGTCGAGCTGTGTA 1191
 Db 1150 ACACAGATTAACATGCTCGCTGTCGAGCGATTTGTCGAGCTGTGTA 1200

RESULT 5
 ECASPC 1415 bp DNA linear BCT 12-SEP-1993
 LOCUS E. coli aspC gene for aspartate aminotransferase.
 DEFINITION X03629
 ACCESSION X03629.1 GI:41010
 VERSION
 KEYWORDS aminotransferase.
 SOURCE Escherichia coli
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Escherichia.
 1 (bases 1 to 1415)
 Fotheringham, I.G., Dacey, S.A., Taylor, P.P., Smith, T.J.,
 Hunter, M.G., Finlay, M.E., Primrose, S.B., Parker, D.M. and
 Edwards, R.M.
 The cloning and sequence analysis of the aspC and tyrB genes from
 Escherichia coli K12. Comparison of the primary structures of the
 aspartate aminotransferase and aromatic aminotransferase of E. coli
 with those of the pig aspartate aminotransferase isoenzymes
 Biochem. J. 234 (3), 593-604 (1986)
 JOURNAL 3521591
 PUBMED
 COMMENT Data kindly reviewed (25-JUN-1986) by M. Hunter.
 FEATURES
 Location/Qualifiers
 1..1415
 /organism="Escherichia coli"
 /mol_type="genomic DNA"
 /db_xref="taxon:562"
 68..73
 /note="pot. promoter signal"
 91..96
 /note="pot. promoter signal"
 104
 /note="mRNA start site"
 127..130
 /note="ribosome binding site"
 138..1328
 /note="unnamed protein product; aspartate aminotransferase (aspC)"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA27279.1"
 /db_xref="GI:41011"
 /db_xref="UniProt/Swiss-Prot:P00509"
 /translation="MFENITTAAPADPILGLADLFRADERPGKINILGIVYKDETKTP
 VLTSVKASQYLLNETTKYLLIDGIPFGRCTOELLFGKGSALINDKRAKTAOTPG
 GTGALRVADFLAKNTSVKRVMSNPSPNHSVNSAGLEVREYAYVDAENHTLDPG
 ALINLSNEAQAGDVVLPFGCCNPTGIDITPLEQWOTLAQLSVEKWLPLDFAYQGF
 RLGEDEGLRPAFAMHKKELIVASSYSKFNGLYNERVAGACTLVAADSETVDRAFOMK
 AAIRANSPNPAGHSAVATILSDNALRAIWEQLDMRQRIORMRQLFVNLTQSKGA
 NRDFSPFIKQNGMFSFSLTKQVLRLEEFVGYVASGRVNVAGMTPDNMAPLCEAI
 VAVL"

ORIGIN

Query Match 100.0%; Score 1191; DB 1; Length 1415;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTGAACAATTAACCCGCGTCTCTGCGACCCCGATTTCTGGGCCCTGCGCGATCTGTTT 60
 Db 138 ATGTTTGAACAATTAACCCGCGTCTCTGCGACCCCGATTTCTGGGCCCTGCGCGATCTGTTT 197
 QY 61 CTTGCGGATGAACGCTCCCGCAAAATTAACCTCGGATTTGCTCTATAAAGATGAGACG 120
 Db 198 CTTGCGGATGAACGCTCCCGCAAAATTAACCTCGGATTTGCTCTATAAAGATGAGACG 257
 QY 121 GGCACAAACCCCGGATCTGACACGCGTGAAGAGGCTGAACAGATATCTGCTCGAAAAATGA 180
 Db 258 GGCACAAACCCCGGATCTGACACGCGTGAAGAGGCTGAACAGATATCTGCTCGAAAAATGA 317
 QY 181 ACACCAAAATTAACCTCGGATTTGACGCGATCCCTGAAATTTGGTCTGCTGCACTCAGGAA 240
 Db 318 ACCACCAAAATTAACCTCGGATTTGACGCGATCCCTGAAATTTGGTCTGCTGCACTCAGGAA 377
 QY 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCAGGACACAGACT 300
 Db 378 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCAGGACACAGACT 437
 QY 301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAAAATACACAGCTT 360
 Db 438 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAAAATACACAGCTT 497
 QY 361 AAGCGTGTGTGGGTGAGCAACCAAGCTGCGCGAACCAATAAGAGCGTCTTTAACTCTGCA 420
 Db 498 AAGCGTGTGTGGGTGAGCAACCAAGCTGCGCGAACCAATAAGAGCGTCTTTAACTCTGCA 557
 QY 421 GGTCTGGAAGTTGCGGAATACGCTTATATGATGCGGAAAAATCACACTTTGACTCTCGAT 480
 Db 558 GGTCTGGAAGTTGCGGAATACGCTTATATGATGCGGAAAAATCACACTTTGACTCTCGAT 617
 QY 481 GCACTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTTCCATGGCTGC 540
 Db 618 GCACTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTTCCATGGCTGC 677
 QY 541 TGCCATAACCCAAACCGGTATCGACCTTACGCTGGAACAATGGAACAACACATGGCACAACTC 600
 Db 678 TGCCATAACCCAAACCGGTATCGACCTTACGCTGGAACAATGGAACAACACATGGCACAACTC 737
 QY 601 TCGGTTGAGAAAGGCTGGTTACGGCTGTTGACTTCGCTTACCAGGGTTTGGCCGCTGGT 660
 Db 738 TCGGTTGAGAAAGGCTGGTTACGGCTGTTGACTTCGCTTACCAGGGTTTGGCCGCTGGT 797
 QY 661 CTGGAAGAAGATGCTGAAGGACTTGGCGGCTTTCGCGGCTATGATATAAGAGCTGATTGTT 720
 Db 798 CTGGAAGAAGATGCTGAAGGACTTGGCGGCTTTCGCGGCTATGATATAAGAGCTGATTGTT 857
 QY 721 GCCAGTTCCTACTCTAAAAAATTTGGCCTGTACACGAGCGTGTGGCGCTTGTACTCTG 780
 Db 858 GCCAGTTCCTACTCTAAAAAATTTGGCCTGTACACGAGCGTGTGGCGCTTGTACTCTG 917
 QY 781 GTTGTGCGCGACGATGAACCCGTTGATTCGCGATTCAGGCAATGAAGCGGCGATTCGC 840
 Db 918 GTTGTGCGCGACGATGAACCCGTTGATTCGCGATTCAGGCAATGAAGCGGCGATTCGC 977
 QY 841 GCTAACTACTCTAAACCCACGACACGCGCTTCTGTTGTTGCCACCACTCTCTGAGCAAC 900
 Db 978 GCTAACTACTCTAAACCCACGACACGCGCTTCTGTTGTTGCCACCACTCTCTGAGCAAC 1037
 QY 901 GATGCGTTACGTCGCGATTTGGGAAACAAGAGCTGACTGATATGCGCCACGGTATTCAGCGT 960
 Db 1038 GATGCGTTACGTCGCGATTTGGGAAACAAGAGCTGACTGATATGCGCCAGCGTATTACCGT 1097
 QY 961 ATGCGTCACTGTTGCTCAATACGCTGAGGAAAAAGGCGCAAAACCGCGACTTCAGCTTT 1020
 Db 1098 ATGCGTCACTGTTGCTCAATACGCTGAGGAAAAAGGCGCAAAACCGCGACTTCAGCTTT 1157
 QY 1021 ATCATCAACAGACGCGCATGTTCTCTTCAGTGGCTGACAAAAAGAAACAAGTGTGCGCT 1080
 Db 1158 ATCATCAACAGACGCGCATGTTCTCTTCAGTGGCTGACAAAAAGAAACAAGTGTGCGCT 1217
 QY 1081 CTGCGGGAAGAGTTTGGCGGTATATGCGGTTCGCTTCTGCTGCGGTAAATGTGCGCGGATG 1140

```
Db 1218 CTCGCGAAGAGTTGGCGTATATGCGTTGCTTCCTGTCGCTAAATGTGGCCGGATG 1277
Qy 1141 ACACAGATAACATGGCTCGCTGTGGGAAGCAATGTGGCAGTGTGTAA 1191
Db 1278 ACACAGATAACATGGCTCGCTGTGGGAAGCAATGTGGCAGTGTGTAA 1328

RESULT 6
LOCUS 108484 1468 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 9 from Patent WO 8700202.
ACCESSION 108484
VERSION 108484.1 GI:588804
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1468)
AUTHORS Edwards,M.R., Taylor,P.P., Hunter,M.G. and Fotheringham,I.G.
TITLE COMPOSITE PLASMIDS FOR AMINO ACID SYNTHESIS
JOURNAL Patent: WO 8700202-A 9 15-JAN-1987;
FEATURES
    source
        1..1468
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 1191; DB 6; Length 1468;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTTGAGAACATTAACCGCGCTCTGCGGACCGCATCTGCGGCCTGGCCGATCTGTTT 60
Db 191 ATGTTTGAGAACATTAACCGCGCTCTGCGGACCGCATCTGCGGCCTGGCCGATCTGTTT 250
Qy 61 CGTGCCGATGAACGTCCTGGCAAAATTAACCTCGGATTTGGTCTATAAAGATGAGCG 120
Db 251 CGTGCCGATGAACGTCCTGGCAAAATTAACCTCGGATTTGGTCTATAAAGATGAGCG 310
Qy 121 GGCAAAACCCGGTACTGACGACGCTGAAAGAGCTGAACAGTATCTGCTCGAAATGAA 180
Db 311 GGCAAAACCCGGTACTGACGACGCTGAAAGAGCTGAACAGTATCTGCTCGAAATGAA 370
Qy 181 ACCACAAATAATTAACCTGGCATTAACCGGATCCCTGAAATTTGGTCTGCTGCACTAGAA 240
Db 371 ACCACAAATAATTAACCTGGCATTAACCGGATCCCTGAAATTTGGTCTGCTGCACTAGAA 430
Qy 241 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGCTGCTCGCAGGCACAGCT 300
Db 431 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGCTGCTCGCAGGCACAGCT 490
Qy 301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGGCGGATTTCTGGCAAAATAATACAGCGTT 360
Db 491 CCGGGGGGCACTGGCGCACTACGCGTGGCTGGCGGATTTCTGGCAAAATAATACAGCGTT 550
Qy 361 AAGCGTGTGGTGAGCAACCCCAAGCTGCCCAACCAATGAGAGCGTCTTAACTCTGCA 420
Db 551 AAGCGTGTGGTGAGCAACCCCAAGCTGCCCAACCAATGAGAGCGTCTTAACTCTGCA 610
Qy 421 GGTCTGGAAGTTCGTGATACGCTTATATGATGCGGAAATCACAATCTTGAATTCGAT 480
Db 611 GGTCTGGAAGTTCGTGATACGCTTATATGATGCGGAAATCACAATCTTGAATTCGAT 670
Qy 481 GCACTGATTAACAGCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTTCAATGCTGC 540
Db 671 GCACTGATTAACAGCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTTCAATGCTGC 730
Qy 541 TGCATAACCCAAACCGGTATCGACCTCAAGCTGGAAACAATGGCAAAACATGGCAAACTC 600
Db 731 TGCATAACCCAAACCGGTATCGACCTCAAGCTGGAAACAATGGCAAAACATGGCAAACTC 790
Qy 601 TCCGTTGAGAAAGGCTGTTACCGCTGTTTGACTTTCGCTTACCAGGGTTTGGCCGTTG 660
```

```
Db 791 TCCGTTGAGAAAGGCTGGTTACCGCTGTTTGACTTCGCTTACAGGGTTTGGCCGTTG 850
Qy 661 CTGGAAGAAGATGCTGAAGGACTGCGGCTTTGCGGGCTATGCATATAAGAGCTGATTGT 720
Db 851 CTGGAAGAAGATGCTGAAGGACTGCGGCTTTGCGGGCTATGCATATAAGAGCTGATTGT 910
Qy 721 GCCAGTCTCTACTCTAAATAAATTTTGGCCTGTACACAGCGCTGTGGCGCTTGTACTCTG 780
Db 911 GCCAGTCTCTACTCTAAATAAATTTTGGCCTGTACACAGCGCTGTGGCGCTTGTACTCTG 970
Qy 781 GTTGTCTGCGCAGTGAACCCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTCGC 840
Db 971 GTTGTCTGCGCAGTGAACCCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTCGC 1030
Qy 841 GCTAACTACTTAACCCACAGCACACGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 900
Db 1031 GCTAACTACTTAACCCACAGCACACGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 1090
Qy 901 GATCGTTACGTCGATTTGGGAACAAGAGCTGACTGATATCGCCAGCGTATTTCAGCGT 960
Db 1091 GATCGTTACGTCGATTTGGGAACAAGAGCTGACTGATATCGCCAGCGTATTTCAGCGT 1150
Qy 961 ATGCGTCAAGTTGTTGTTCAATACGCTGCAGGAAAAAGCGCAAAACCGCGACTTCAGCTTT 1020
Db 1151 ATGCGTCAAGTTGTTGTTCAATACGCTGCAGGAAAAAGCGCAAAACCGCGACTTCAGCTTT 1210
Qy 1021 ATCATCAAAACAGAACGGCATGTTTCCTTCAGTGGCTGCACAAAAGAACAGTGTGCGT 1080
Db 1211 ATCATCAAAACAGAACGGCATGTTTCCTTCAGTGGCTGCACAAAAGAACAGTGTGCGT 1270
Qy 1081 CTGCGGAGAGTTTGGCGTATATCGGTTGCTTCTGTCGCTAAATGTCGCCGGGATG 1140
Db 1271 CTGCGGAGAGTTTGGCGTATATCGGTTGCTTCTGTCGCTAAATGTCGCCGGGATG 1330
Qy 1141 ACACAGATAAATGCTCCGCTGTCGAAAGCGATTGTGGCAGTGTCTGTA 1191
Db 1331 ACACAGATAAATGCTCCGCTGTCGAAAGCGATTGTGGCAGTGTCTGTA 1381

RESULT 7
LOCUS 108487 3659 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 15 from Patent WO 8700202.
ACCESSION 108487
VERSION 108487.1 GI:588807
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3659)
AUTHORS Edwards,M.R., Taylor,P.P., Hunter,M.G. and Fotheringham,I.G.
TITLE COMPOSITE PLASMIDS FOR AMINO ACID SYNTHESIS
JOURNAL Patent: WO 8700202-A 15 15-JAN-1987;
FEATURES
    source
        1..3659
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 1191; DB 6; Length 3659;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTTGAGAACATTAACCGCGCTCTGCGGACCGCATTCCTGGGCTGCGCGATCTGTTT 60
Db 2344 ATGTTTGAGAACATTAACCGCGCTCTGCGGACCGCATTCCTGGGCTGCGCGATCTGTTT 2403
Qy 61 CTGCGCGATGAACGTCCTGGGCAAAATTAACCTCGGATTTGGTCTATAAAGATGAGCG 120
Db 2404 CTGCGCGATGAACGTCCTGGGCAAAATTAACCTCGGATTTGGTCTATAAAGATGAGCG 2463
Qy 121 GSCAAAACCCCGTACTGACGCGTGAAAGAGGCTGAACAGTATCTCTCGAAATGAA 180
```

Db 2464 GGCAAAACCCGGTACTGACAGCGTGAAAAGGCTGAAACGATATCTGCTCGAAATGAA 2523
QY 181 ACCACCAAAATTTACCTCGGCATTTGACGGCATCCCTGGAATTTGGTGGCTGCCTCAGGAA 240
Db 2524 ACCACCAAAATTTACCTCGGCATTTGACGGCATCCCTGGAATTTGGTGGCTGCCTCAGGAA 2583
QY 241 CTGCTGTTGGTAAAGTAGCGCCTGATCAATGACAAACGTCGTCGCGGACACAGACT 300
Db 2584 CTGCTGTTGGTAAAGTAGCGCCTGATCAATGACAAACGTCGTCGCGGACACAGACT 2643
QY 301 CCGGGGGGCACTGGCGCACTACGGTGGCTGCGGATTTCTTGCAAAAAATACAGGTT 360
Db 2644 CCGGGGGGCACTGGCGCACTACGGTGGCTGCGGATTTCTTGCAAAAAATACAGGTT 2703
QY 361 AAGCGTGTGGGTGAGCAACCAAGCTGGCCGAACCAATAAGACGCTCTTAACTCTGCA 420
Db 2704 AAGCGTGTGGGTGAGCAACCAAGCTGGCCGAACCAATAAGACGCTCTTAACTCTGCA 2763
QY 421 GGTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAATCACACTCTTGACTTCGAT 480
Db 2764 GGTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAATCACACTCTTGACTTCGAT 2823
QY 481 GCACGTATTAAAGCCTGAATGAAGCTCAGGCTGGCGACGATGCTGTGTTCCATGCTGC 540
Db 2824 GCACGTATTAAAGCCTGAATGAAGCTCAGGCTGGCGACGATGCTGTGTTCCATGCTGC 2883
QY 541 TGGCATAACCCCAACCGGTATCGACCTTACGCTCGGAACATGGAACACATGGGCACAACTC 600
Db 2884 TGGCATAACCCCAACCGGTATCGACCTTACGCTCGGAACATGGAACACATGGGCACAACTC 2943
QY 601 TCGGTTGAGAAAGGCTGGTTACCGCTGTTTGACTTCGCTTACCAGGCTTTTGCCCGTGGT 660
Db 2944 TCGGTTGAGAAAGGCTGGTTACCGCTGTTTGACTTCGCTTACCAGGCTTTTGCCCGTGGT 3003
QY 661 CTGGAAGAAGATCTGAAGACTCGCGCTTTCGCGCTATGCTAAGAGCTGATGTT 720
Db 3004 CTGGAAGAAGATCTGAAGACTCGCGCTTTCGCGCTATGCTAAGAGCTGATGTT 3063
QY 721 GCCAGTTCCTACTCTAANAATTTGGCCTGTACAAGAGGCTGTGGCGCTTGACTCTG 780
Db 3064 GCCAGTTCCTACTCTAANAATTTGGCCTGTACAAGAGGCTGTGGCGCTTGACTCTG 3123
QY 781 GTTGCTGCGCACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGGCAATTCG 840
Db 3124 GTTGCTGCGCACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGGCAATTCG 3183
QY 841 GCTAACTACTTAACCCACAGCACACGGCGCTTCTGTTGGCCACCATCTTGACCAAC 900
Db 3184 GCTAACTACTTAACCCACAGCACACGGCGCTTCTGTTGGCCACCATCTTGACCAAC 3243
QY 901 GATGCGTTACGTCGGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCT 960
Db 3244 GATGCGTTACGTCGGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCT 3303
QY 961 ATGCGTCAGTTGTTGCTCAATACGCTGAGGAAAAAGGCGCAACCGCACTTCAGCTTT 1020
Db 3304 ATGCGTCAGTTGTTGCTCAATACGCTGAGGAAAAAGGCGCAACCGCACTTCAGCTTT 3363
QY 1021 ATCATCAAAACAGAACGGCATTTGTTCTCTTCAGTGGCTTGACAAAGAACAGTGTGCGT 1080
Db 3364 ATCATCAAAACAGAACGGCATTTGTTCTCTTCAGTGGCTTGACAAAGAACAGTGTGCGT 3423
QY 1081 CTGCGCGAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGCGTAAATGTGCGCGGATG 1140
Db 3424 CTGCGCGAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGCGTAAATGTGCGCGGATG 3483
QY 1141 ACACCAAGATAACATGCTCGCTGTCGGAAGCGATTTGTCAGTGTCTGTA 1191
Db 3484 ACACCAAGATAACATGCTCGCTGTCGGAAGCGATTTGTCAGTGTCTGTA 3534

D90731/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

D90731
Escherichia coli K12 genomic DNA. (21.2 - 21.5 min). linear BCT 25-DEC-2002
D90731 AB001340
D90731.1 GI:1651452
Complete and shotgun sequencing; ycbB; aspC; ompF; tolF; cmlB; coa; cry; asnS; tss; pncB; pepN; ycbE; yzeB.
Escherichia coli K12
Escherichia coli K12
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
1
Oshima, T., Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiiuchi, T., Ikemoto, K., Inada, T., Itoh, T., Kajihara, M., Kanai, K., Kashimoto, K., Kimura, S., Kitagawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishio, Y., Oshima, T., Saito, N., Yamamoto, Y., Yano, M. and Horiuchi, T.
A 718-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 12.7-28.0 min region on the linkage map
DNA Res. 3 (3), 137-155 (1996)
8905232
2
Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiuchi, T., Ikemoto, K., Inada, T., Isono, S., Itoh, T., Kanai, K., Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M., Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishio, Y., Oshima, T., Saito, N., Sampei, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C., Yamamoto, Y. and Yano, M.
The systematic sequencing of the Escherichia coli genome in Japan
Unpublished
3 (bases 1 to 11470)
Mori, H.
Direct Submission
Submitted (29-JUL-1996) Hirotsada Mori, NARA Institute of Science and Technology, Res. & Edu. Center for Genetic Info.; 8916-5 Takayama, Ikoma, Nara 630-01, Japan
(E-mail:hmori@gctc.aist-nara.ac.jp, Tel:81-7437-2-5660, Fax:81-7437-2-5669)
Collaboration Information:
Project:
The Japan E.coli genome DNA sequencing project
Group:
The Japan E.coli genome DNA sequencing group
Members: (1995.4 - 1996.3)
Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiuchi, T., Ikemoto, K., Inada, T., Isono, K., Isono, S., Itoh, T., Kanai, K., Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M., Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishio, Y., Oshima, T., Saito, N., Sampei, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C., Yamamoto, Y. and Yano, M.
Headed by:
Name: Takashi Horiuchi
Address: National Institute of Basic Biology, Okazaki, 444, Japan
E-mail: kishori@nibb.ac.jp
Information Operator:
Name: Hirotsada Mori
Address: NARA Institute of Science and Technology, Ikoma, 630-01, Japan
E-mail: hmori@gctc.aist-nara.ac.jp
URL:
The Japan E. coli genome database
http://bsw3.aist-nara.ac.jp.
FEATURES
Location/Qualifiers
1..11470
/organism="Escherichia coli K12"
/mol_type="genomic DNA"
/strain="K12"
/db_xref="taxon:83333"
/map="21.2-21.5 min"

|||||
2682 ACCACAAAATTACCTCGCATTTGACGGCATCCCTGAATTTGGTCGCTGCACTCAGGAA 2623
Qy
241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAAGTGTCTCGCAGGACAGACT 300
Db
2622 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAAGTGTCTCGCAGGACAGACT 2563
Qy
301 CCGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAATACAGCGTT 360
Db
2562 CCGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAATACAGCGTT 2503
Qy
361 AAGCGTGTGGGTAGCAACCCAAAGTGGCGGCAACCAATAGAGCGTCTTTAACTCTGCA 420
Db
2502 AAGCGTGTGGGTAGCAACCCAAAGTGGCGGCAACCAATAGAGCGTCTTTAACTCTGCA 2443
Qy
421 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
Db
2442 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 2383
Qy
481 GCACTGATTAAAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTCTGTTCCATGGCTGC 540
Db
2382 GCACTGATTAAAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTCTGTTCCATGGCTGC 2323
Qy
541 TGGCATAACCAACCGGTATCGACCCCTACGCTGGAACAAATGGCAACACTGGCACAACTC 600
Db
2322 TGGCATAACCAACCGGTATCGACCCCTACGCTGGAACAAATGGCAACACTGGCACAACTC 2263
Qy
601 TCGGTTGAGAAAGCTGGTACGCTGTTTGACTTCGCTTACGAGGTGTTTGGCGGTGGT 660
Db
2262 TCGGTTGAGAAAGCTGGTACGCTGTTTGACTTCGCTTACGAGGTGTTTGGCGGTGGT 2203
Qy
661 CTGGAAGAATGCTGAAGGACTGCGCGCTTTTCGCGCTATGATGAAGAGCTGATTGTT 720
Db
2202 CTGGAAGAATGCTGAAGGACTGCGCGCTTTTCGCGCTATGATGAAGAGCTGATTGTT 2143
Qy
721 GCCAGTTTCTTACTCTAAAACTTTGGCTCTGTAACAAGAGCGTGTGGCGCTTGTACTCTG 780
Db
2142 GCCAGTTTCTTACTCTAAAACTTTGGCTCTGTAACAAGAGCGTGTGGCGCTTGTACTCTG 2083
Qy
781 GTTGCTGCCGACAGTGAAAACCGTTGATCGCGCATTCAGCCAAATGAAGCGGCGATTGCG 840
Db
2082 GTTGCTGCCGACAGTGAAAACCGTTGATCGCGCATTCAGCCAAATGAAGCGGCGATTGCG 2023
Qy
841 GCTAACTTACTCTAACCCACAGACACGCGCTTCTGTTGTCGCCACCATCTCGAGCAAC 900
Db
2022 GCTAACTTACTCTAACCCACAGACACGCGCTTCTGTTGTCGCCACCATCTCGAGCAAC 1963
Qy
901 GATGCGTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGGCGCCAGCGTATTGAGCGT 960
Db
1962 GATGCGTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGGCGCCAGCGTATTGAGCGT 1903
Qy
961 ATGCGTCAGTTGTTGCTCAATACGCTCCAGGAAAAGCGCAACCCGCACTTCAGCTTT 1020
Db
1902 ATGCGTCAGTTGTTGCTCAATACGCTCCAGGAAAAGCGCAACCCGCACTTCAGCTTT 1843
Qy
1021 ATCATCAACAGAACGCGATGTTCTCTCAGTGGCCCTGACAAAGAACAAAGTGTCTCGT 1080
Db
1842 ATCATCAACAGAACGCGATGTTCTCTCAGTGGCCCTGACAAAGAACAAAGTGTCTCGT 1783
Qy
1081 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTCGCTGCTGCTGCTGCTGCTG 1140
Db
1782 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTCGCTGCTGCTGCTGCTGCTG 1723
Qy
1141 ACACGAGATACATGGCTCCGCTGTGCGAAGCGATTTGTGCGAGTGTCTGTAA 1191
Db
1722 ACACGAGATACATGGCTCCGCTGTGCGAAGCGATTTGTGCGAGTGTCTGTAA 1672

RESULT 9
AX370270/c
LOCUS
DEFINITION
ACCESSION

AX370270
Sequence 87 from Patent WO0170776.
AX370270

14759 bp
DNA
linear
PAT 16-FEB-2002

AX370270.1 GI:18857433
Escherichia coli
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1
Levy, S.B., Barbosa, T.M. and Alekshun, M.N.
Nimr compositions and their methods of use
Patent: WO 0170776-A 87 27-SEP-2001;
TRUSTEES OF TUFTS COLLEGE (US)
FEATURES
source
1. 14759
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
ORIGIN
Query Match 100.0%; Score 1191; DB 6; Length 14759;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTTTGAGAAATATACCGCGCTCTGCGACCCCGATTTCTGGGCTTGGCGCATCTGTTT 60
Db 2755 ATGTTTGAGAAATATACCGCGCTCTGCGACCCCGATTTCTGGGCTTGGCGCATCTGTTT 2696
Qy 61 CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTCTATATAAGATGAGACG 120
Db 2695 CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTCTATATAAGATGAGACG 2636
Qy 121 GGCAAAACCCCGGTACTGACGCGTGAACCAAGGCTGAAACAGTATCTGCTCGAAAATGAA 180
Db 2635 GGCAAAACCCCGGTACTGACGCGTGAACCAAGGCTGAAACAGTATCTGCTCGAAAATGAA 2576
Qy 181 ACCACAAAATTAACCTCGGATTTGACGCGATCCCTGAAATTTGGTCTGCTGCACTCAGAA 240
Db 2575 ACCACAAAATTAACCTCGGATTTGACGCGATCCCTGAAATTTGGTCTGCTGCACTCAGAA 2516
Qy 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCCTCGCAGGACACACT 300
Db 2515 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCCTCGCAGGACACACT 2456
Qy 301 CCGGGGGCACTGGCGCACTTACGCGTGGCTGCGGATTTCTTGGCAAAAATACAGCGTT 360
Db 2455 CCGGGGGCACTGGCGCACTTACGCGTGGCTGCGGATTTCTTGGCAAAAATACAGCGTT 2396
Qy 361 AAGCGTGTGGGTAGCAACCCAAAGTGGCGGCAACCAATAGAGCGTCTTTAACTCTGCA 420
Db 2395 AAGCGTGTGGGTAGCAACCCAAAGTGGCGGCAACCAATAGAGCGTCTTTAACTCTGCA 2336
Qy 421 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
Db 2335 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 2276
Qy 481 GCACTGATTAAAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTCTGTTCCATGGCTGC 540
Db 2275 GCACTGATTAAAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTCTGTTCCATGGCTGC 2216
Qy 541 TGGCATAACCAACCGGTATCGACCCCTACGCTGGAACAAATGGCAACACTGGCAACAACTC 600
Db 2215 TGGCATAACCAACCGGTATCGACCCCTACGCTGGAACAAATGGCAACACTGGCAACAACTC 2156
Qy 601 TCGGTTGAGAAAGCTGGTTTACCGCTGTTTGACTTCGCTTACCGGGTGGTGGCGGTGGT 660
Db 2155 TCGGTTGAGAAAGCTGGTTTACCGCTGTTTGACTTCGCTTACCGGGTGGTGGCGGTGGT 2096
Qy 661 CTGGAAGAATGCTGAAGGACTGCGCGCTTTTCGCGCTATGATGAAGAGCTGATTGTT 720
Db 2095 CTGGAAGAATGCTGAAGGACTGCGCGCTTTTCGCGCTATGATGAAGAGCTGATTGTT 2036
Qy 721 GCCAGTTTCTTACTCTAAAACTTTGGCTCTGTAACAAGAGCGTGTGGCGCTTGTACTCTG 780
Db 2035 GCCAGTTTCTTACTCTAAAACTTTGGCTCTGTAACAAGAGCGTGTGGCGCTTGTACTCTG 1976


```

/protein_id="BAA35662.1"
/db_xref="GI:4062488"
/translation="MSLPHLSADARNHLHAAQLGNKPRRRASLEDIPATISRMILL
QIDTINVARSPYLVLFPSRLGNYPQAWLDESARGELMEVYHAECFMPSPDFRLIRH
RLMAPGKMKYKADAMQWHEAEIAQLIHIHQKGPVRSADFPHKPSGASGWWEKPH
KRLHEGLTAGKVMVIERNNFORVYDLTHRWKPDWDDERDLVSQTEAEIIMLDSARS
LGFRLQWLLADYRLRPAALAAWEAREAEQQOIIAIVHEKLGKLNWLDHLLPLLERAL
AGKLTATHSAVLSPFDPVMDRKRARQLFDFSVRLCEYTPAPKROYGYVLPILLHRGQ
LVGRMDAKHROQTGILEVLSLQEGIKPTTTLQKGLRQAITDFANWQAQATRVTLGCC
PQGLFTDCRTGWEIDPA"
3027..3209
/notes="unnamed protein product; ORF_ID:o218#3"
/codon_start=1
/transl_table=11
/protein_id="BAA35663.1"
/db_xref="GI:4062489"
/translation="MDHRLLEIITACVNCGLWYNQEKQELICKLNLNLAFFLRDGPV
LLETEARVUTADESKS"
3206..3952
/genes="kdsB"
3206..3952
/genes="kdsB"
/notes="ORF_ID:o218#4
similar to PIR Accession Number A26322"
/codon_start=1
/transl_table=11
/product="3-deoxy-manno-oculosonate cytidylyltransferase
(EC 2.7.7.38)"
/protein_id="BAA35664.1"
/db_xref="GI:1651444"
/notes="unnamed protein product; ORF_ID:o218#5"
/codon_start=1
/transl_table=11
/protein_id="BAA35665.1"
/db_xref="GI:4062490"
/translation="MQRLAEKLSHLGKLSRIECVNEKADTALWALYDSQGNPMLP
ARSFSPGKARQLAWTKMLARSCTGVMTPIYGVMTHEHPGPDVLLLRMRGVSVEA
PARTPERWEQLQDOIYALLAHWRDSCGCVADVNTQENFWPSYRQHVLEVLTWLN
QENNTGLTWQDKLILRTSECLPALPEGNDNCVLJHGFCLRSLMKDSRDLQWLV
TGLRLMWAPEYELFLRLMNSLAEDLLMSYLQRPVAESFIWRRLYLVLWDEAVLUN
GGRFSRNFDLASKSLPLMLA"
complement(4976..5755)
/genes="ybcB"
complement(4976..5755)
/genes="ybcB"
/notes="ORF_ID:o218#6
similar to SwissProt Accession Number P36565"
/codon_start=1
/transl_table=11
/product="Hypothetical protein in kdsB-mukF intergenic
region"
/protein_id="BAA35666.1"
/db_xref="GI:4062491"
/translation="MLFTLKVKVIGNNMLPLPLMLLIIGAGLALLWFSRFQKTGKIFIS
IGWLALLLSLOQVADRLRLPIESTYPTWNSOKVDYIVVLGGYTNWNPQWAPSSNLI
NNSLPRNLRSIRLWRNPGSKLIFTGGVAKTNTVSTAEVGARVAQSISGVPRQIITLD
LPKDTSEAAAVKQAGDAPFLLVTSASHLPRAMIFFQQEGLNPLPAPANQLAIDSPL
NPWERAIPSPVLMHSDRVGYETLGRWQWLKGSSEGPQES"
5891..6676
/genes="smtA"
5891..6676
/genes="smtA"
/notes="ORF_ID:o218#7
similar to SwissProt Accession Number P36566"
/codon_start=1
/transl_table=11
/product="SmtA protein."

```

```

/protein_id="BAA35667.1"
/db_xref="GI:1651445"
/translation="MODRNFDDIAEKFSSRIYVTTKQLROAILWODLDRVLAEMGPQ
KURVLDAGGEGQTAIKMAERGHQVILCDLSAOMIDRAKQAAEAKGVSDNQPTHCAA
QDVASHLETVPDLILFHAVLEWADPRSVLOTLSVLPRGCVLSLFPYNAHGLLMHNM
VAGNFYVQAGMPKKKRTLSPDYPHDPAPQVILWLEAGAQIMGKTGVRVHFHYLREK
HQORDCYEALLLELEYTRCYRQEPYITLGRYIHVTARKPQSKDKV"
6673..7995
/genes="kicB"
/notes="synonym: mukF"
6673..7995
/genes="kicB"
/notes="ORF_ID:o218#8"

```

Query Match	100.0%	Score 1191	DB 1	Length 20604
Best Local Similarity	100.0%	Pred No. 0		
Matches 1191	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	ATGTTTGAGAACATTACCGCCGCTCTCGCGACCCGATTCGGCCCTGGCCGATCTGTTT	60	
DB	18063	ATGTTTGAGAACATTACCGCCGCTCTCGCGACCCGATTCGGCCCTGGCCGATCTGTTT	18004	
QY	61	CGTGCCGATGAACGTCGCCGCAAAATTAACCTCGGATTCGGTCTGTATATAAGATGACACG	120	
DB	18003	CGTGCCGATGAACGTCGCCGCAAAATTAACCTCGGATTCGGTCTGTATATAAGATGACACG	17944	
QY	121	GGCAAAACCCCGTACTGACGCGTGAAGAAAGCTGAACAGTATCTGCTCGAAATGAA	180	
DB	17943	GGCAAAACCCCGTACTGACGCGTGAAGAAAGCTGAACAGTATCTGCTCGAAATGAA	17884	
QY	181	ACCACCAAAATTAACCTCGGCATTGACGGCATCCCTGAAATTTGGTCTGCTGCACCTCAGGAA	240	
DB	17883	ACCACCAAAATTAACCTCGGCATTGACGGCATCCCTGAAATTTGGTCTGCTGCACCTCAGGAA	17824	
QY	241	CTGCTCTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACAGCTGCTGCCACGACACACT	300	
DB	17823	CTGCTCTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACAGCTGCTGCCACGACACACT	17764	
QY	301	CGGGGGGCGACCTGGCGCCTACGCGTGGCTGCCGATTTCCCTGCAAAAATACCAGCGTT	360	
DB	17763	CGGGGGGCGACCTGGCGCCTACGCGTGGCTGCCGATTTCCCTGCAAAAATACCAGCGTT	17704	
QY	361	AAGCGTGTGTGGGTGAGCAACCCAAAGCTGGCGCAACCAATAAGAGCGCTTTAACTCTGCA	420	
DB	17703	AAGCGTGTGTGGGTGAGCAACCCAAAGCTGGCGCAACCAATAAGAGCGCTTTAACTCTGCA	17644	
QY	421	GGTCTGGAAGTTGCTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT	480	
DB	17643	GGTCTGGAAGTTGCTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT	17584	
QY	481	GCACGTATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGGCTGC	540	
DB	17583	GCACGTATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGGCTGC	17524	
QY	541	TGCATAACCCAAACCGGTATCGACCCCTAGCTCGAAACAATGGCAACAACATGGCACAACTC	600	
DB	17523	TGCATAACCCAAACCGGTATCGACCCCTAGCTCGAAACAATGGCAACAACATGGCACAACTC	17464	
QY	601	TCCGTTGAGAAAGCGTGGTTACCGCTGTTTGAATTCGCTTACAGGGTTTTCGCCCGTGT	660	
DB	17463	TCCGTTGAGAAAGCGTGGTTACCGCTGTTTGAATTCGCTTACAGGGTTTTCGCCCGTGT	17404	
QY	661	CTGGAGAGAGATGCTGAAGACTGCGGCTTTTCGGCTTATGATATAAGAGCTGATTGTT	720	
DB	17403	CTGGAGAGAGATGCTGAAGACTGCGGCTTTTCGGCTTATGATATAAGAGCTGATTGTT	17344	
QY	721	GCCAGTTCCTACTCTAAAACTTTTGGCTGTACAACGAGCGGTGTGGCGCTTGTACTCTG	780	
DB	17343	GCCAGTTCCTACTCTAAAACTTTTGGCTGTACAACGAGCGGTGTGGCGCTTGTACTCTG	17284	
QY	781	GTTGCTGCCGACAGTGAACCCGTTGATCGCGCATTCAGCCAAATGAAGAGCGGATTCGC	840	
DB	17283	GTTGCTGCCGACAGTGAACCCGTTGATCGCGCATTCAGCCAAATGAAGAGCGGATTCGC	17224	

Qy	841	GCTAACTACTCTAACCCACGACACAGCGGCTTCGTGTTGTCACCATCTCGACCAAC	900
Db	17223	GCTAACTACTCTAACCCACGACACAGCGGCTTCGTGTTGTCACCATCTCGACCAAC	17164
Qy	901	GATGCGTTACGTCGCAATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACAGCGT	960
Db	17163	GATGCGTTACGTCGCAATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACAGCGT	17104
Qy	961	ATGCGTCAGTTGTTCTCAATACGCTGCAGGAAAAAGGCGCAAAACCGCACTTTCAGCTTTT	1020
Db	17103	ATGCGTCAGTTGTTCTCAATACGCTGCAGGAAAAAGGCGCAAAACCGCACTTTCAGCTTTT	17044
Qy	1021	ATCATCAAAACAGAACGGCATGTTCTCTTCAGTGGGCTGACAAAAGAACAGTGTGCGT	1080
Db	17043	ATCATCAAAACAGAACGGCATGTTCTCTTCAGTGGGCTGACAAAAGAACAGTGTGCGT	16984
Qy	1081	CTGCGGAAGAGTTTGGCGTATATGCGTTGCTTCTGCTGCGTAAATGTGGCCGGGATG	1140
Db	16983	CTGCGGAAGAGTTTGGCGTATATGCGTTGCTTCTGCTGCGTAAATGTGGCCGGGATG	16924
Qy	1141	ACACCAGATAACATGCTCGCTGTGCGAAGCGAATTTGGCAGTGTCTGTAA	1191
Db	16923	ACACCAGATAACATGCTCGCTGTGCGAAGCGAATTTGGCAGTGTCTGTAA	16873

RESULT 11
U00096 09/c
WPCOMMENT

Sequence split into 47 fragments LOCUS U00096 Accession U00096

Fragment Name	Begin	End
U00096_00	1	110000
U00096_01	100001	210000
U00096_02	200001	310000
U00096_03	300001	410000
U00096_04	400001	510000
U00096_05	500001	610000
U00096_06	600001	710000
U00096_07	700001	810000
U00096_08	800001	910000
U00096_09	900001	1010000
U00096_10	1000001	1110000
U00096_11	1100001	1210000
U00096_12	1200001	1310000
U00096_13	1300001	1410000
U00096_14	1400001	1510000
U00096_15	1500001	1610000
U00096_16	1600001	1710000
U00096_17	1700001	1810000
U00096_18	1800001	1910000
U00096_19	1900001	2010000
U00096_20	2000001	2110000
U00096_21	2100001	2210000
U00096_22	2200001	2310000
U00096_23	2300001	2410000
U00096_24	2400001	2510000
U00096_25	2500001	2610000
U00096_26	2600001	2710000
U00096_27	2700001	2810000
U00096_28	2800001	2910000
U00096_29	2900001	3010000
U00096_30	3000001	3110000
U00096_31	3100001	3210000
U00096_32	3200001	3310000
U00096_33	3300001	3410000
U00096_34	3400001	3510000
U00096_35	3500001	3610000
U00096_36	3600001	3710000
U00096_37	3700001	3810000
U00096_38	3800001	3910000
U00096_39	3900001	4010000
U00096_40	4000001	4110000
U00096_41	4100001	4210000

U00096_42	4200001	4310000
U00096_43	4300001	4410000
U00096_44	4400001	4510000
U00096_45	4500001	4610000
U00096_46	4600001	4639675
Continuation (10 of 47) of U00096 from base 900001 (U00096 Escherichia coli K-12		
Query Match 100.0%; Score 1191; DB 1; Length 110000;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	ATGTTTGAGAACATTACCGCGCTCTCTGCCGACCCGATTTCTGGGCTTGGCCGATCTGTTT 60
Db	84932	ATGTTTGAGAACATTACCGCGCTCTCTGCCGACCCGATTTCTGGGCTTGGCCGATCTGTTT 84873
Qy	61	CGTCCGATGAACGTCCTCCGGCAAAATTAACCTCGGGATTTGGTGTCTATAAAGATGAGACG 120
Db	84872	CGTCCGATGAACGTCCTCCGGCAAAATTAACCTCGGGATTTGGTGTCTATAAAGATGAGACG 84813
Qy	121	GGCAAAACCCCGGTACTGACCGAGCTGAAAAGGCTGAACTATCTGCTCGAAAAATGAA 180
Db	84812	GGCAAAACCCCGGTACTGACCGAGCTGAAAAGGCTGAACTATCTGCTCGAAAAATGAA 84753
Qy	181	ACCACCAAAATTTACCTCGGCATTGACGGCATCCCTGAAATTTGGTTCGCTGCATCAGGAA 240
Db	84752	ACCACCAAAATTTACCTCGGCATTGACGGCATCCCTGAAATTTGGTTCGCTGCATCAGGAA 84693
Qy	241	CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCGACGGCACAGACT 300
Db	84692	CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCGACGGCACAGACT 84633
Qy	301	CGGCGGGGCACTGGCGCACTACGCGTGGCTGCCGATTTCTGGCAAAAATACACAGCGTT 360
Db	84632	CGGCGGGGCACTGGCGCACTACGCGTGGCTGCCGATTTCTGGCAAAAATACACAGCGTT 84573
Qy	361	AAGCGTGTGCGGTGAGCAACCCAAAGCTGGCCGCAACCAATAGAGCGCTCTTTAACTCTGCA 420
Db	84572	AAGCGTGTGCGGTGAGCAACCCAAAGCTGGCCGCAACCAATAGAGCGCTCTTTAACTCTGCA 84513
Qy	421	GCTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAAAATCACACTCTTTGACTTCGAT 480
Db	84512	GCTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAAAATCACACTCTTTGACTTCGAT 84453
Qy	481	GCACTGATTAACAGCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGCTGCTGC 540
Db	84452	GCACTGATTAACAGCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGCTGCTGC 84393
Qy	541	TGCCATAACCCAAACCGGTATCGACCCCTACGCTGGAACAAATGGCAAACTGGCAACAATC 600
Db	84392	TGCCATAACCCAAACCGGTATCGACCCCTACGCTGGAACAAATGGCAAACTGGCAACAATC 84333
Qy	601	TCCGTTGAGAAAGGCTGTTTACCGCTGTTTGACTTTCGCTTACAGGGTTTTGCCCGTGGT 660
Db	84332	TCCGTTGAGAAAGGCTGTTTACCGCTGTTTGACTTTCGCTTACAGGGTTTTGCCCGTGGT 84273
Qy	661	CTGGAAGAAGATGCTGAAGGACTCGCGCTTTTGGCGCTATGCAATAAAGAGCTGATTGTT 720
Db	84272	CTGGAAGAAGATGCTGAAGGACTCGCGCTTTTGGCGCTATGCAATAAAGAGCTGATTGTT 84213
Qy	721	GCCAGTTCCTACTCTATAAAATTTTGGCTGTGACACGAGCGGTGTTGGCGTTGTACTCTG 780
Db	84212	GCCAGTTCCTACTCTATAAAATTTTGGCTGTGACACGAGCGGTGTTGGCGTTGTACTCTG 84153
Qy	781	GTTGCTGCCGACAGTGAACCGTTGATTCGCGATTCAGCCAAATGAAGCGGATTCGC 840
Db	84152	GTTGCTGCCGACAGTGAACCGTTGATTCGCGATTCAGCCAAATGAAGCGGATTCGC 84093
Qy	841	GCTAACTACTCTAACCCACGACACAGCGGCTTCTGTTGTTGCCACCACTCTGAGCAAC 900
Db	84092	GCTAACTACTCTAACCCACGACACAGCGGCTTCTGTTGTTGCCACCACTCTGAGCAAC 84033
Qy	901	GATGCGTTACGTCGCAATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960

Db	84032	GATCGGTTACGTGCGATTTCGGGAACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT	83973
Qy	961	ATGCGTCAGTGTGTTTCGTCATACCTGACGAGAAAGCGCAAAACCGCGACTTCAGCTTT	1020
Db	83972	ATGCGTCAGTGTGTTTCGTCATACCTGACGAGAAAGCGCAAAACCGCGACTTCAGCTTT	83913
Qy	1021	ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCCTGACAAAGAACAAAGTGTGCGT	1080
Db	83912	ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCCTGACAAAGAACAAAGTGTGCGT	83853
Qy	1081	CTGGCGGAAGATTGGCGTATATGCGGTTGCTTCTGTCGCGTAATGTGGCGGGATG	1140
Db	83852	CTGGCGGAAGATTGGCGTATATGCGGTTGCTTCTGTCGCGTAATGTGGCGGGATG	83793
Qy	1141	ACACCAAGTAACATGGCTCCGCTGTGGAAGCGATTGTGGCAGTGTGTAA	1191
Db	83792	ACACCAAGTAACATGGCTCCGCTGTGGAAGCGATTGTGGCAGTGTGTAA	83742
RESULT 12			
AE005174_11/c			
WPCOMMENT			
Sequence split into 56 fragments LOCUS AE005174 Accession AE005174			
Fragment Name Begin End			
AE005174_00	1	110000	
AE005174_01	100001	210000	
AE005174_02	200001	310000	
AE005174_03	300001	410000	
AE005174_04	400001	510000	
AE005174_05	500001	610000	
AE005174_06	600001	710000	
AE005174_07	700001	810000	
AE005174_08	800001	910000	
AE005174_09	900001	1010000	
AE005174_10	1000001	1110000	
AE005174_11	1100001	1210000	
AE005174_12	1200001	1310000	
AE005174_13	1300001	1410000	
AE005174_14	1400001	1510000	
AE005174_15	1500001	1610000	
AE005174_16	1600001	1710000	
AE005174_17	1700001	1810000	
AE005174_18	1800001	1910000	
AE005174_19	1900001	2010000	
AE005174_20	2000001	2110000	
AE005174_21	2100001	2210000	
AE005174_22	2200001	2310000	
AE005174_23	2300001	2410000	
AE005174_24	2400001	2510000	
AE005174_25	2500001	2610000	
AE005174_26	2600001	2710000	
AE005174_27	2700001	2810000	
AE005174_28	2800001	2910000	
AE005174_29	2900001	3010000	
AE005174_30	3000001	3110000	
AE005174_31	3100001	3210000	
AE005174_32	3200001	3310000	
AE005174_33	3300001	3410000	
AE005174_34	3400001	3510000	
AE005174_35	3500001	3610000	
AE005174_36	3600001	3710000	
AE005174_37	3700001	3810000	
AE005174_38	3800001	3910000	
AE005174_39	3900001	4010000	
AE005174_40	4000001	4110000	
AE005174_41	4100001	4210000	
AE005174_42	4200001	4310000	
AE005174_43	4300001	4410000	
AE005174_44	4400001	4510000	
AE005174_45	4500001	4610000	
AE005174_46	4600001	4710000	
AE005174_47	4700001	4810000	
AE005174_48	4800001	4910000	

AE005174_49	4900001	5010000	
AE005174_50	5000001	5110000	
AE005174_51	5100001	5210000	
AE005174_52	5200001	5310000	
AE005174_53	5300001	5410000	
AE005174_54	5400001	5510000	
AE005174_55	5500001	5528445	
Continuation (12 of 56) of AE005174 from base 1100001 (AE005174 Escherichia coli O157:H7)			
Query Match 98.13%; Score 1170.2; DB 1; Length 110000;			
Best Local Similarity 98.9%; Pred. No. 0;			
Matches 1178; Conservative 0; Mismatches 13; Indels 0; Gaps 0;			
Qy	1	ATGTTTGGAGAACATTACCGCCGCTCTCGCGACCCCGATTCTGGGCCCTGGCCGATCTCTTT	60
Db	104403	ATGTTTGGAGAACATTACCGCCGCTCTCGCGACCCCGATTCTGGGCCCTGGCCGATCTCTTT	104344
Qy	61	CTGCGCGATGAACGTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG	120
Db	104343	CTGCGCGATGAACGTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG	104284
Qy	121	GGCAAAACCCCGTACTGACCGCATCCCTGAAATTTGGTCTGCTGCACCTCAGGAA	180
Db	104283	GGCAAAACCCCGTACTGACCGCATCCCTGAAATTTGGTCTGCTGCACCTCAGGAA	104224
Qy	181	ACCACCAAAATTTACCTCGGCATTGACCGCATCCCTGAAATTTGGTCTGCTGCACCTCAGGAA	240
Db	104223	ACCACCAAAATTTACCTCGGCATTGACCGCATCCCTGAAATTTGGTCTGCTGCACCTCAGGAA	104164
Qy	241	CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACCGTCTGCGACGGACAGACT	300
Db	104163	CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACCGTCTGCGACGGACAGACT	104104
Qy	301	CCGGGGGCACTGGCGCATCTACGGTGGCTGCGGATTTCTTGGCAAAATATCCAGGTT	360
Db	104103	CCGGGGGCACTGGCGCATCTACGGTGGCTGCGGATTTCTTGGCAAAATATCCAGGTT	104044
Qy	361	AAGCGTGTGGGTGAGCAACCCAAAGCTGGCCCAACCATTAAGAGCGTCTTTAACTCTGCA	420
Db	104043	AAGCGTGTGGGTGAGCAACCCAAAGCTGGCCCAACCATTAAGAGCGTCTTTAACTCTGCA	103984
Qy	421	GGTCTGGAAGTTCGTGAATAACGCTTTATATGATGCGGAAATCACACTTTGACTTCGAT	480
Db	103983	GATCTGGAAGTTCGTGAATAACGCTTTATATGATGCGGAAATCACACTTTGACTTCGAT	103924
Qy	481	GCACTGATTAAACAGCTGTAATGAACTCAGGCTGGCGACGACTGCTGTTCCATGGCTGC	540
Db	103923	GCACTGATTAAACAGCTGTAATGAACTCAGGCTGGCGACGACTGCTGTTCCATGGCTGC	103864
Qy	541	TGCAATAACCCACCGTATCGACCTAGCTGGAAACAAATGGCAAAACACTGGCACAACTC	600
Db	103863	TGCAATAACCCACCGTATCGACCTAGCTGGAAACAAATGGCAAAACACTGGCACAACTC	103804
Qy	601	TCGGTTGAGAAAGGCTGGTTTACCGCTGTTTGAATTCGCTTTACCAGGGTTTGGCCGTTGGT	660
Db	103803	TCGGTTGAGAAAGGCTGGTTTACCGCTGTTTGAATTCGCTTTACCAGGGTTTGGCCGTTGGT	103744
Qy	661	CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGGGGCTATGCAATAAAGAGCTGATTGTT	720
Db	103743	CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGGGGCTATGCAATAAAGAGCTGATTGTT	103684
Qy	721	GCCAGTTCTCTACTTAAACCTTTTGGCCCTGTCAACAGAGCGTGTGGCGCTTGTACTCTG	780
Db	103683	GCCAGTTCTCTACTTAAACCTTTTGGCCCTGTCAACAGAGCGTGTGGCGCTTGTACTCTG	103624
Qy	781	GTTGCTGCCGACAGTGAACACCGTTGATTCGGCATTCAGCCAAATGAAGCGGCGATTTCG	840
Db	103623	GTTGCTGCCGACAGTGAACACCGTTGATTCGGCATTCAGCCAAATGAAGCGGCGATTTCG	103564
Qy	841	GCTAACTACTTAACCCACAGACACGGCGCTTCTGTTGTTGCGCACCATCTCTGAGCAAC	900
Db	103563	GCTAACTACTTAACCCACAGACACGGCGCTTCTGTTGTTGCGCACCATCTCTGAGCAAC	103504

Db 3563 GCTAACTACTCTAAACCCACGACACACGGCGCTTCTGTGTTGCCACCATCTCTGAGCAAC 3504

Qy 901 GATGCGTTACGTGCTGATTTGGAAACAAGAGCTGATGATATGCCCGCAGCGTATTACGCGT 960

Db 3503 GATGCGTTACGTGCGATTTGGAAACAAGAGCTGATGATATGCCCGCAGCGTATTACGCGT 3444

Qy 961 ATGCGTCAGTTGTCGTCATACGCTGCAGGAAAGCGCAAAACCGGACTTTCAGCTTTT 1020

Db 3443 ATGCGTCAGTTGTCGTCATACGCTGCAGGAAAGCGCAAAACCGGACTTTCAGCTTTT 3384

Qy 1021 ATCATCAACAGACGCGCATGTTCTCCTTCAGTGGCTGACAAAGAACAAAGTCTGCGT 1080

Db 3383 ATCATCAACAGACGCGCATGTTCTCCTTCAGTGGCTGACAAAGAACAAAGTCTGCGT 3324

Qy 1081 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGCGTAAATGTCGCGGATG 1140

Db 3323 CTGCGCAAGAGTTTGGCGTATGCTGTTGCTTCTGCTGCGTAAACGTTGGCGGATG 3264

Qy 1141 ACACCGATTAACATGGCTCCGCTGTGCGAAGCGAATTTGCGAGTGTGTAA 1191

Db 3263 ACACCGATTAACATGGCTCCGCTGTGCGAAGCGAATTTGCGAGTGTGTAA 3213

RESULT 14

BA000007_11/c

WPCOMMENT

Sequence split into 55 fragments LOCUS BA000007 Accession BA000007

Fragment Name	Begin	End
BA000007_00	1	110000
BA000007_01	100001	210000
BA000007_02	200001	310000
BA000007_03	300001	410000
BA000007_04	400001	510000
BA000007_05	500001	610000
BA000007_06	600001	710000
BA000007_07	700001	810000
BA000007_08	800001	910000
BA000007_09	900001	1010000
BA000007_10	1000001	1110000
BA000007_11	1100001	1210000
BA000007_12	1200001	1310000
BA000007_13	1300001	1410000
BA000007_14	1400001	1510000
BA000007_15	1500001	1610000
BA000007_16	1600001	1710000
BA000007_17	1700001	1810000
BA000007_18	1800001	1910000
BA000007_19	1900001	2010000
BA000007_20	2000001	2110000
BA000007_21	2100001	2210000
BA000007_22	2200001	2310000
BA000007_23	2300001	2410000
BA000007_24	2400001	2510000
BA000007_25	2500001	2610000
BA000007_26	2600001	2710000
BA000007_27	2700001	2810000
BA000007_28	2800001	2910000
BA000007_29	2900001	3010000
BA000007_30	3000001	3110000
BA000007_31	3100001	3210000
BA000007_32	3200001	3310000
BA000007_33	3300001	3410000
BA000007_34	3400001	3510000
BA000007_35	3500001	3610000
BA000007_36	3600001	3710000
BA000007_37	3700001	3810000
BA000007_38	3800001	3910000
BA000007_39	3900001	4010000
BA000007_40	4000001	4110000
BA000007_41	4100001	4210000
BA000007_42	4200001	4310000
BA000007_43	4300001	4410000
BA000007_44	4400001	4510000

BA000007_45	4500001	4610000
BA000007_46	4600001	4710000
BA000007_47	4700001	4810000
BA000007_48	4800001	4910000
BA000007_49	4900001	5010000
BA000007_50	5000001	5110000
BA000007_51	5100001	5210000
BA000007_52	5200001	5310000
BA000007_53	5300001	5410000
BA000007_54	5400001	5498450

Continuation (12 of 55) of BA0000007 from base 1100001 (BA0000007 Escherichia coli)

Query Match	98.3%;	Score 1170.2;	DB 1;	Length 110000;
Best Local Similarity	98.9%;	Pred. No. 0;		
Matches 1178;	Conservative	0;	Mismatches 13;	Indels 0; Gaps 0;

QY	1	ATGTTTGGAGAACATTACCGCCGCTCTCTGCGGACCCGATTCCTGGCCCTGCGCCGATCTGTTT	60
DB	15180	ATGTTTGGAGAACATTACCGCGCTCTCTGCGGACCCGATTCCTGGCCCTGCGCCGATCTGTTT	1512
QY	61	CGTGCCGATGAACGTCCTCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG	120
DB	15120	CGTGCCGATGAACGTCCTCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG	1506
QY	121	GGCAAAACCCGGTACTGACGAGCTGAAAGGCTGAACAGTATCTGCTCGAAATGAA	180
DB	15060	GGCAAAACCCGGTACTGACGAGCTGAAAGGCTGAACAGTATCTGCTCGAAATGAA	1500
QY	181	ACCACCAAAATTTACCTCGGCATTGACGCATCCCTGAAATTTGGTCGCTGCACCTCAGGAA	240
DB	15000	ACCACCAAAATTTACCTCGGCATTGACGCATCCCTGAAATTTGGTCGCTGCACCTCAGGAA	1494
QY	241	CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGCTGCTCGCACGGCACAGACT	300
DB	14940	CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGCTGCTCGCACGGCACAGACT	1488
QY	301	CCGGGGGCGACTGCGGCACCTACGCGTGGCTGCGGATTTCTTGGCAAAAAATACACGGCTT	360
DB	14880	CCGGGTGGCACTGGCGCACTACGCATAGCTGCGGATTTCTTGGCAAAAAATACACGGCTT	1482
QY	361	AAGCGTGTGGGTGAGCAACCCAGCTGCGCCGAACCATGAAGAGCGCTTTAACTCTGCA	420
DB	14820	AAGCGAGTGTGGGTGAGCAACCCAGCTGCGCGAACCATAAAGAGCGCTTTAACTCTGCA	1476
QY	421	GGTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAAAATCACACTCTTGACTTCGAT	480
DB	14760	GATCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAAAATCACACCCCTTGACTTCGAT	1470
QY	481	GCACTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAAGTGTGTTTCCATGGCTGC	540
DB	14700	GCACTGATTAAACAGCCTGAACGAAGCTCAGGCTGGCGACGTAAGTGTGTTTCCATGGCTGC	1464
QY	541	TGCCATAACCCAAACCGGTATCGACCTACGCTGGAACAATGGCAACACATGGGCACAACTC	600
DB	14640	TGCCACAACCCAAACCGGTATCGACCCCTACGCTGGAACAATGGCAGACACATGGGCACAACTC	1458
QY	601	TCCGTTGAGAAAGGCTGGTTTACCGCTGTTTGACTTCGCTTACCAGGGTTTTCGCCCGTGGT	660
DB	14580	TCCGTTGAGAAAGGCTGGTTTACCGCTGTTTGACTTCGCTTACCAGGGTTTTCGCCCGTGGT	1452
QY	661	CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGGCTATGCATAAAGAGCTGATTGTT	720
DB	14520	CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGGCTATGCATAAAGAGCTGATTGTT	1446
QY	721	GCCAGTTCCTACTCTAAAAAATTTTGGCTGTACAACAGCGGTGTTTGGCGCTTGTACTCTG	780
DB	14460	GCCAGTTCCTACTCTAAAAAATTTTGGCTGTACAACAGCGGTGTTTGGCGCTTGTACTCTG	1440
QY	781	GTTGCTGCCGACAGTGAACCCGTTGATCGCGATTTCAGCCAAATGAAGCGGCGATTTCGC	840
DB	14400	GTTGCTGCCGACAGTGAACCCGTTGATCGCGATTTCAGCCAAATGAAGCGGCGATTTCGC	1434
QY	841	GCTAACTACTCTAAACCCACGACACACGCGGCTTCTGTTGTTGCCACCATCTCTGAGCAAC	900

||||| 14340 GCTAACTACTCTAACCACCAGCACAGCGGCTTCTGTTGTTGCCACCATTCTCTGAGCAAC 14281
QY 901 GATGCGTTACGTCGGATTTGGGAAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCGT 960
Db 14280 GATGCGTTACGTCGGATTTGGGAAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCGT 14221
QY 961 ATGGGTCAGTTGTCGTCATACGCTGCAGGAAAAGGCGCAACCGGACTTTCAGCTTT 1020
Db 14220 ATGGGTCAGTTGTCGTCATACGCTGCAGGAAAAGGCGCAACCGGACTTTCAGCTTT 14161
QY 1021 ATCATCAAAACAGAACGGCATGTTCTCTTCAGTGGCTGACAAAAAGAAACAAGTGTGCGT 1080
Db 14160 ATCATCAAAACAGAACGGCATGTTCTCTTCAGTGGCTGACAAAAAGAAACAAGTGTGCGT 14101
QY 1081 CTGCGGGAAGAGTTTGGCGTATATGCGGTGCTTCTGCTGCGTAAATGTGGCCGGGATG 1140
Db 14100 CTGCGGGAAGAGTTTGGCGTATATGCTGTTGCTTCTGCTGCGTAAACGTTGGCCGGGATG 14041
QY 1141 ACACCAAGATAACATGCTCGCTGTCGGAAGCGATTGTGGCAGTGTCTGTAA 1191
Db 14040 ACACCAAGATAACATGCTCGCTGTCGGAAGCGATTGTGGCAGTGTCTGTAA 13990

RESULT 15

AE005674_09/c

WPCOMMENT

Sequence split into 46 fragments LOCUS AE005674 Accession AE005674

Fragment Name	Begin	End
AE005674_00	1	110000
AE005674_01	100001	210000
AE005674_02	200001	310000
AE005674_03	300001	410000
AE005674_04	400001	510000
AE005674_05	500001	610000
AE005674_06	600001	710000
AE005674_07	700001	810000
AE005674_08	800001	910000
AE005674_09	900001	1010000
AE005674_10	1000001	1110000
AE005674_11	1100001	1210000
AE005674_12	1200001	1310000
AE005674_13	1300001	1410000
AE005674_14	1400001	1510000
AE005674_15	1500001	1610000
AE005674_16	1600001	1710000
AE005674_17	1700001	1810000
AE005674_18	1800001	1910000
AE005674_19	1900001	2010000
AE005674_20	2000001	2110000
AE005674_21	2100001	2210000
AE005674_22	2200001	2310000
AE005674_23	2300001	2410000
AE005674_24	2400001	2510000
AE005674_25	2500001	2610000
AE005674_26	2600001	2710000
AE005674_27	2700001	2810000
AE005674_28	2800001	2910000
AE005674_29	2900001	3010000
AE005674_30	3000001	3110000
AE005674_31	3100001	3210000
AE005674_32	3200001	3310000
AE005674_33	3300001	3410000
AE005674_34	3400001	3510000
AE005674_35	3500001	3610000
AE005674_36	3600001	3710000
AE005674_37	3700001	3810000
AE005674_38	3800001	3910000
AE005674_39	3900001	4010000
AE005674_40	4000001	4110000
AE005674_41	4100001	4210000
AE005674_42	4200001	4310000
AE005674_43	4300001	4410000

AE005674_44 4400001 4510000
AE005674_45 4500001 4607203
Continuation (10 of 46) of AE005674 from base 900001 (AB005674 Shigella flexneri 2a str
Query Match 98.1%; Score 1168.6; DB 1; Length 110000;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1177; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 ATGTTTGAGAACTTACCGCGCTCTGCCGACCCGATTCTGGCGCTGCGCGCATCTGTTT 60
Db 70933 ATGTTTGAGAACTTACCGCGCTCTGCCGACCCGATTCTGGCGCTGCGCGCATCTGTTT 70874
QY 61 CGTGGCGATGAACGTCGCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 120
Db 70873 CGTGGCGATGAACGTCGCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 70814
QY 121 GGCACAAACCCGGTACTGACAGCGTGAAAAAGGTGAAACAGTATCTGCTCGAAAAATGAA 180
Db 70813 GGCACAAACCCGGTACTGACAGCGTGAAAAAGGTGAAACAGTATCTGCTCGAAAAATGAA 70754
QY 181 ACCACCAAAATTTACTCGGCATTGACGCGATCCCTGAAATTTGGTGTGCTGCACTCAGGAA 240
Db 70753 ACCACCAAAATTTACTCGGCATTGACGCGATCCCTGAAATTTGGTGTGCTGCACTCAGGAA 70694
QY 241 CTGCTGTTTGGTAAAGGTAGCGCTGATCAATGACAAACGCTGCTCGCACGGCACAGACT 300
Db 70693 CTGCTGTTTGGTAAAGGTAGCGCTGATCAATGACAAACGCTGCTCGCACGGCACAGACT 70634
QY 301 CCGGGGGCAGCTGGCGCACTACGCGTGGCTGCCGATTTCTCGGCAAAAAAATACCAGCGTT 360
Db 70633 CCGGGTGGCAGCTGGCGCACTACGCGTGGCTGCCGATTTCTCGGCAAAAAAATACCAGCGTT 70574
QY 361 AAGCGTGTGGGTGAGCAACCCAGCTGGCCGACCAATAGAGCGTCTTTAACTCTGCA 420
Db 70573 AAGCGTGTGGGTGAGCAACCCAGCTGGCCGACCAATAGAGCGTCTTTAACTCTGCA 70514
QY 421 GCTCTGGAAGTTCTGTGAAATACGCTTATTATGATGCGAAAAATCACACTCTTGACTTCGAT 480
Db 70513 GCTCTGGAAGTTCTGTGAAATACGCTTATTATGATGCGAAAAATCACACTCTTGACTTCGAT 70454
QY 481 GCACTGATTAACAGCTCTGAACGAGCTCAGGCTGGCGAGTAGTGTCTCCATGCGCTGC 540
Db 70453 GCACTGATTAACAGCTCTGAACGAGCTCAGGCTGGCGAGTAGTGTCTCCATGCGCTGC 70394
QY 541 TGCCATACCCAAACCGGTATCGACCCCTAGCTGGAAACANTGGCAACACTGCGCACACTC 600
Db 70393 TGCCATACCCAAACCGGTATCGACCCCTAGCTGGAAACANTGGCAACACTGCGCACACTC 70334
QY 601 TCCGTTGAGAAAGGCTGGTTACCGCTGTTTGACTTTCGCTTACACAGGGTTTTGCCCCGTGGT 660
Db 70333 TCCGTTGAGAAAGGCTGGTTACCGCTGTTTGACTTTCGCTTACACAGGGTTTTGCCCCGTGGT 70274
QY 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGGCTATGCATAAAGAGCTGATTGTT 720
Db 70273 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGGCTATGCATAAAGAGCTGATTGTT 70214
QY 721 GCCAGTTCTACTCTAAAAAATTTTGGCGCTGTACAAAGAGCGTGTGGCGCTTGACTCTCG 780
Db 70213 GCCAGTTCTACTCTAAAAAATTTTGGCGCTGTACAAAGAGCGTGTGGCGCTTGACTCTCG 70154
QY 781 GTTGTGCGGACAGTGAAACCGTTGATCGCGATTTCAGCCAAATGAAAGCGGCGATTTCG 840
Db 70153 GTTGTGCGGACAGTGAAACCGTTGATCGCGATTTCAGCCAAATGAAAGCGGCGATTTCG 70094
QY 841 GCTAACTACTCTAACCCACAGCACACGGCGCTTCTGTTGTTGCTGCCACCACTCTTGAGCAAC 900
Db 70093 GCTAACTACTCTAACCCACAGCACACGGCGCTTCTGTTGTTGCTGCCACCACTCTTGAGCAAC 70034
QY 901 GATGCGTTACGTGCGATTGGGAAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCGT 960
Db 70033 GATGCGTTACGTGCGATTGGGAAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCGT 69974
QY 961 ATGCGTCAAGTTGTTTCGTCAATACGCTGACAGGAAAAAGGCGCAACCGCGACTTCAGCTTT 1020

[illegible]

Search completed: March 15, 2006, 15:50:55
Job time : 6324 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 17:52:30 ; Search time 189 Seconds
(without alignments)

920.603 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045

Sequence: 1 MFENITAAPADPILGLADLF.....VAGMTDPNMAPLCEAIVAVL 396

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_21:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*
- 9: geneseqp2005s:*

*Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2045	100.0	396	1	AAP71678
2	2045	100.0	396	2	AAW69553
3	2045	100.0	396	7	ADB83285
4	2045	100.0	396	9	ADN18132
5	2045	100.0	396	9	ADW95404
6	1931	94.4	402	7	ABO65556
7	1682	82.2	397	6	ABM69171
8	1654	80.9	396	8	ADS42695
9	1318	64.4	322	8	ADN17578
10	1317	64.4	423	6	ABP80003
11	1311	64.1	397	5	AAU73004
12	1310	64.1	397	8	ADP08222
13	1005.5	49.2	394	8	ADS22791
14	996	48.7	396	8	ADN24844
15	996	48.7	398	8	ADN22085
16	982	48.0	397	8	ADN25771
17	982	48.0	413	7	ABO80253
18	975	47.7	397	8	ADS26795
19	975	47.7	397	8	ADS27175
20	975	47.7	398	8	ADS26427
21	947	46.3	398	8	ADS24992
22	924	45.2	395	8	ADN26805
23	917	44.8	395	8	ADN26569
24	917	44.8	420	8	ADY07798
					Plant ful

25	914	44.7	420	8	ADY07427	Ady07427 Plant ful
26	913	44.6	423	8	ADX94364	Adx94364 Plant ful
27	911.5	44.6	395	8	ADS24690	AdS24690 Bacterial
28	911	44.5	417	8	ADY12894	Ady12894 Plant ful
29	909.5	44.5	543	7	ABO80332	ABO80332 Pseudomon
30	909	44.4	400	8	ADS28454	AdS28454 Bacterial
31	908	44.4	233	7	ADF07692	AdF07692 Bacterial
32	893	43.7	396	8	ADN26006	Adn26006 Bacterial
33	884.5	43.3	396	8	ADS42724	AdS42724 Bacterial
34	876	42.8	410	3	AAG12869	Aag12869 Arabidops
35	876	42.8	449	3	AAG12868	Aag12868 Arabidops
36	876	42.8	453	3	AAG12867	Aag12867 Arabidops
37	875.5	42.8	396	8	ADN17634	Adn17634 Bacterial
38	874	42.7	410	3	AAG43050	Aag43050 Arabidops
39	874	42.7	410	3	AAG50737	Aag50737 Arabidops
40	874	42.7	419	3	AAG50729	Aag50729 Arabidops
41	874	42.7	449	3	AAG43049	Aag43049 Arabidops
42	874	42.7	449	3	AAG50736	Aag50736 Arabidops
43	874	42.7	453	3	AAG43048	Aag43048 Arabidops
44	874	42.7	453	3	AAG50735	Aag50735 Arabidops
45	874	42.7	458	3	AAG50728	Aag50728 Arabidops

ALIGNMENTS

RESULT 1
AAP71678
ID AAP71678 standard; protein; 396 AA.
XX
AC AAP71678;
XX
DT 01-JAN-1980 (first entry)
XX
DE aspc gene product from plasmid pME219.
XX
KW alpha-amylase; feedback inhibition; amino acid synthesis;
KW composite plasmid.
XX
OS Bacillus licheniformis.
XX
FN W08700202-A.
XX
PD 15-JAN-1987.
XX
PF 24-JUN-1986; 86WO-US001353.
XX
PR 24-JUN-1985; 85US-00747732.
XX
PA (NUTR-) NUTRASWEET CO.
XX
PI Edwards MR, Taylor PP, Hunter MG, Fotheringh IG;
XX
DR WPI; 1987-021998/03.
XX
N-PSDB; AAN71109.
PT Composite plasmids contg. multiple genes in transcriptional units -
PT useful for prodn. of aminoacid(s), esp. L-phenylalanine and l-tyrosine.
XX
Disclousure; Page 38; 57pp; English.
XX
PS App80003 N. gonorr
XX
PS Aau73004 Neisseria
XX
CC Adp08222 Neisseria
XX
CC AdS22791 Bacterial
XX
CC Adn24844 Bacterial
XX
CC Adn22085 Bacterial
XX
CC Adn25771 Bacterial
XX
CC Abo80253 Pseudomon
XX
CC Aag26795 Bacterial
XX
CC Aag27175 Bacterial
XX
CC Aag26427 Bacterial
XX
CC Aag24992 Bacterial
XX
CC Adn26805 Bacterial
XX
CC Adn26569 Bacterial
XX
SQ Sequence 396 AA;
Query Match 100.0%; Score 2045; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No 2e-197;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFADLERPGKINIGVYKDETKPTVLTSTVKAEQYLLENE 60
|||||

Db 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETKTPVLTSVKKAQYLLENE 60
Qy 61 TTKNYLGIDGIPFGRCCTOELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120
Db 61 TTKNYLGIDGIPFGRCCTOELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120
Qy 121 KRVVSNPSPNHNKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLPFGC 180
Db 121 KRVVSNPSPNHNKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLPFGC 180
Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKGWLPDFDFAVQGFARGLEDEAGLRAFAAMHKLIV 240
Db 181 CHNPTGIDPTLEQWQTLAQLSVEKGWLPDFDFAVQGFARGLEDEAGLRAFAAMHKLIV 240
Qy 241 ASSYSKNFGLYNERVAGCTILVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYSKNFGLYNERVAGCTILVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Qy 301 DALRAIWEQELTDMRQRIQMRQLFVNTLQEKGNRDFSFIIKQNGMFSFGLTKQVLR 360
Db 301 DALRAIWEQELTDMRQRIQMRQLFVNTLQEKGNRDFSFIIKQNGMFSFGLTKQVLR 360
Qy 361 LREBFGVYAVASGRVNVAGTMDNMAPLCEAIVAVL 396
Db 361 LREBFGVYAVASGRVNVAGTMDNMAPLCEAIVAVL 396

RESULT 2
AAW69553
ID AAW69553 standard; protein; 396 AA.
XX AC AAW69553;
XX DT 13-OCT-1998 (first entry)
XX DE Escherichia coli aspC protein.
XX KW Brevibacterium lactofermentum; lysC; L-lysine; coryneform bacterium;
XX KW aspartokinase; feedback inhibition; dihydropicolinate reductase;
XX KW diaminopimelate decarboxylase; aspartate aminotransferase.
XX OS Escherichia coli.
XX PN EP854189-A2.
XX PD 22-JUL-1998.
XX PP 05-DEC-1997; 97EP-00121443.
XX PR 05-DEC-1996; 96JP-00325659.
XX PA (AJIN) AJINOMOTO CO INC.
XX PI Araki M, Sugimoto M, Yoshihara Y, Nakamatsu T;
XX DR WPI; 1998-379060/33.
XX DR N-PSDB; AAV40259.
XX PT Recombinant DNA autonomously replicable in coryneform bacteria - used to
PT produce L-lysine, codes for e.g. aspartokinase, dihydropicolinate
PT reductase and synthase and di.amino-pimelate decarboxylase.
XX PS Claim 6; Page 38-39; 59pp; English.
XX CC The present invention describes a recombinant DNA autonomously replicable
CC in cells of coryneform bacteria (CB), comprising a DNA sequence coding
CC for an aspartokinase (AK) in which feedback inhibition by L-lysine and L-
CC threonine is desensitised, a DNA sequence coding for a
CC dihydropicolinate reductase (DHPR), a DNA sequence coding for
CC dihydropicolinate synthase (DHPS), a DNA sequence coding for
CC diaminopimelate decarboxylase (DAMD) and a DNA sequence coding for
CC aspartate aminotransferase (AAT). The present sequence represents aspC
CC from Escherichia coli. The DNA and related products from the present

CC invention, can be used for improving L-lysine productivity by CB. The L-
CC lysine produced can be used as a fodder additive
SQ Sequence 396 AA;
Query Match 100.0%; Score 2045; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 2e-197;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETKTPVLTSVKKAQYLLENE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETKTPVLTSVKKAQYLLENE 60
Qy 61 TTKNYLGIDGIPFGRCCTOELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120
Db 61 TTKNYLGIDGIPFGRCCTOELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120
Qy 121 KRVVSNPSPNHNKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLPFGC 180
Db 121 KRVVSNPSPNHNKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLPFGC 180
Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKGWLPDFDFAVQGFARGLEDEAGLRAFAAMHKLIV 240
Db 181 CHNPTGIDPTLEQWQTLAQLSVEKGWLPDFDFAVQGFARGLEDEAGLRAFAAMHKLIV 240
Qy 241 ASSYSKNFGLYNERVAGCTILVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYSKNFGLYNERVAGCTILVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Qy 301 DALRAIWEQELTDMRQRIQMRQLFVNTLQEKGNRDFSFIIKQNGMFSFGLTKQVLR 360
Db 301 DALRAIWEQELTDMRQRIQMRQLFVNTLQEKGNRDFSFIIKQNGMFSFGLTKQVLR 360
Qy 361 LREBFGVYAVASGRVNVAGTMDNMAPLCEAIVAVL 396
Db 361 LREBFGVYAVASGRVNVAGTMDNMAPLCEAIVAVL 396
RESULT 3
ADB83285
ID ADB83285 standard; protein; 396 AA.
XX AC ADB83285;
XX DT 04-DEC-2003 (first entry)
XX DE Escherichia coli aspartate aminotransferase.
XX KW aspartate aminotransferase; threonine; fermentation; enzyme.
XX OS Escherichia coli.
XX PN WO2003072786-A1.
XX PD 04-SEP-2003.
XX PP 25-FEB-2003; 2003WO-JP002067.
XX PR 27-FEB-2002; 2002RU-00104983.
XX PA (AJIN) AJINOMOTO CO INC.
XX PI Akhverdian VZ, Savrasova EA, Kaplan AM, Lobanov AO, Kozlov YI;
XX DR WPI; 2003-721777/68.
XX DR N-PSDB; ADB83284.
XX PT Industrial production of L-threonine by fermentation using Escherichia
PT modified to enhance aspartate aminotransferase activity, with improved
PT productivity.
XX PS Claim 7; Page 20-21; 26pp; Japanese.

CC The invention relates to a bacterium belonging to the genus *Escherichia*
CC which is modified to enhance aspartate aminotransferase activity for the
CC production of L-threonine. The bacterium is used for the industrial
CC production of L-threonine by fermentation. This sequence corresponds to
CC the *E. coli* aspartate aminotransferase protein.
XX
SQ Sequence 396 AA;

Query Match 100.0%; Score 2045; DB 7; Length 396;
Best Local Similarity 100.0%; Pred. No. 2e-197;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINLGIGYKDETKTPTVLTSVKKAQVYLLNE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINLGIGYKDETKTPTVLTSVKKAQVYLLNE 60
Qy 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTAQTPGGTGALRVAADFLAKNTSV 120
Db 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTAQTPGGTGALRVAADFLAKNTSV 120
Qy 121 KRVVNSPNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLPFGC 180
Db 121 KRVVNSPNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLPFGC 180
Qy 181 CHNPTGIDPTLEQWOTLAQSVKGMWLPDFAYQGFARGLEDAEGLRFAAMHKLIV 240
Db 181 CHNPTGIDPTLEQWOTLAQSVKGMWLPDFAYQGFARGLEDAEGLRFAAMHKLIV 240
Qy 241 ASSYKKNFGLYNERVAGCTLVAAADSETVDRAFQSKAAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYKKNFGLYNERVAGCTLVAAADSETVDRAFQSKAAIRANYSNPPAHGASVVATILSN 300
Qy 301 DALRAIWEQELTDMRQIRQMRLQVNTLQKGNRDFSFIIKQNGMFSFSGLTKEQVLR 360
Db 301 DALRAIWEQELTDMRQIRQMRLQVNTLQKGNRDFSFIIKQNGMFSFSGLTKEQVLR 360
Qy 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 4
ADN18132
ID ADN18132 standard; protein; 396 AA.

AC ADN18132;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #785.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polypeptide.

OS Bacteria.
XX
XX US2003233675-A1.
XX
PD 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX

PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 785; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

Sequence 396 AA;

Query Match 100.0%; Score 2045; DB 8; Length 396;
Best Local Similarity 100.0%; Pred. No. 2e-197;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINLGIGYKDETKTPTVLTSVKKAQVYLLNE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINLGIGYKDETKTPTVLTSVKKAQVYLLNE 60
Qy 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTAQTPGGTGALRVAADFLAKNTSV 120
Db 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTAQTPGGTGALRVAADFLAKNTSV 120
Qy 121 KRVVNSPNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLPFGC 180
Db 121 KRVVNSPNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLPFGC 180
Qy 181 CHNPTGIDPTLEQWOTLAQSVKGMWLPDFAYQGFARGLEDAEGLRFAAMHKLIV 240
Db 181 CHNPTGIDPTLEQWOTLAQSVKGMWLPDFAYQGFARGLEDAEGLRFAAMHKLIV 240
Qy 241 ASSYKKNFGLYNERVAGCTLVAAADSETVDRAFQSKAAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYKKNFGLYNERVAGCTLVAAADSETVDRAFQSKAAIRANYSNPPAHGASVVATILSN 300
Qy 301 DALRAIWEQELTDMRQIRQMRLQVNTLQKGNRDFSFIIKQNGMFSFSGLTKEQVLR 360
Db 301 DALRAIWEQELTDMRQIRQMRLQVNTLQKGNRDFSFIIKQNGMFSFSGLTKEQVLR 360
Qy 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 5

2

QY 241 ASSYSKNFGLYNERVAGCTLVAAADSETVDRAFQKAAIRANYSNPPAHGASVATILSN 300
 DB 247 ASSYSKNFGLYNERVAGCTLVAAADQETVDRAFQKSVIRANYSNPPAHGASVATILSN 306
 QY 301 DALRAIWEQELTDMRQRIORMQLFVNTLOEKGANDRDFSIKQNGMFSFSLGTLKEQVLR 360
 DB 307 DALRAIWEQELTDMRQRIORMQLFVNTLOEKGASRDFSIQNGMFSFSLGTLKEQVLR 366
 QY 361 LREEFGVYAVASGRVNVAGMTDDNAPLCEAIVAVL 396
 DB 367 LREEFAIYAVASGRINVAGMTDDNAPLCEAIVAVL 402

RESULT 7

ID ABM69171 standard; protein; 397 AA.
 XX
 AC ABM69171;

20-NOV-2003 (first entry)

Photorhabdus luminescens protein sequence #2268.

XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough.

XX Photorhabdus luminescens.

OS WO200294867-A2.

PN 28-NOV-2002.

PD 07-FEB-2002; 2002WO-IB003040.

XX 07-FEB-2001; 2001FR-00001659.

XX (INST) INST PASTEUR.

PA (CNRS) CNRS CENT NAT RECH SCI.

XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;

XX WPI; 2003-148459/14.

XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX Claim 2; SEQ ID NO 2268; 1205pp; French.

XX The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens proteins

XX

SQ Sequence 397 AA;

Query Match 82.2%; Score 1682; DB 6; Length 397;

Best Local Similarity 79.8%; Pred. No. 1e-160;
 Matches 316; Conservative 38; Mismatches 42; Indels 0; Gaps 0;

QY 1 MFENITAAPADPILGLADLFRADERPGKINLIGIVYKDETGTPTVLTSVKKAEOYLLENE 60
 DB 1 MFEKITAAPADPILGLADSFSDPTNKINLIGIVYKDETGTPTVLTSVKKAEOYLLENE 60
 QY 61 TTKNYLGIDGIPFGRCCTOELLFGKGSALINDKRAATQTPGCTGALRVAADFALAKNTSV 120
 DB 61 TTKNYLPISGLAEFGKRVTOELLFGKHDPVWTDKARTAQSPGCTGALRIAADFIAKQTNA 120
 QY 121 KRVVSNPNSNHNKSVFNSAGLEVPEYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
 DB 121 KRVWISNPTWPNHNKVFNSAAGLEVCEYKYDYDAEKHALNFMEDLASLSEAQAQGDVVLFHGC 180
 QY 181 CHNPTGIDPTLEQMOTLAQLSVEKGWLPFLFAYOGFARGLEDEAGLRAFAAMHKLIV 240
 DB 181 CHNPTGIDPTPAQWAKLAEMSAEKWLPFDFAYOGFARGLEDEAGLRIFAKNHNLIV 240
 QY 241 ASSYSKNFGLYNERVAGCTLVAAADSETVDRAFQKAAIRANYSNPPAHGASVATILSN 300
 DB 241 ASSYSKNFGLYNERVAGCTIVASDSDTAEKAFSQAKAIRANYSNPPAHGASIVTILSN 300
 QY 301 DALRAIWEQELTDMRQRIORMQLFVNTLOEKGANDRDFSIKQNGMFSFSLGTLKEQVLR 360
 DB 301 EDLKAAMEQELTTKRRIQRMQLFVNTLOEKGAKQDFSIISQNGMFSFSLGTLKEQVER 360
 QY 361 LREEFGVYAVASGRVNVAGMTDDNAPLCEAIVAVL 396
 DB 361 LREEFGIYAVSSGRINVAGLTLENAPLCEAIVAVL 396

RESULT 8

ID ADS42695 standard; protein; 396 AA.
 XX
 AC ADS42695;

XX 02-DEC-2004 (first entry)

XX Bacterial polypeptide #21125.

XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.

OS Bacteria.

PN US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

DR

XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX
PS Claim 1; SEQ ID NO 21125; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 396 AA;

Query Match 80.9%; Score 1654; DB 8; Length 396;
Best Local Similarity 78.3%; Pred. No. 7.1e-158;
Matches 310; Conservative 34; Mismatches 55; Indels 0; Gaps 0;
QY 1 MFENITAAPADPILGLADLFRADERPGKINLGIGYKDETKTPTLTSVKKAEQVLLNE 60
DB 1 MFEKITAAPADPILGLADSFADPRENKINLGIGYKDETKTPTLTVTKKAEFLLENE 60
QY 61 TTKNYLGIIDGPEFCRGTQELLFGKGSALINDKARTATQGTGTCALRVADFLAKNTSV 120
DB 61 TTKNYLAISGLPEFGRVQTQLLFGNTSTIITDKGARTVQSPGTCALRTAADFIAKQTN 120
QY 121 KRVVSNPSPNHNKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLPFGC 180
DB 121 KRVWISNPTWPNHKGVSAGLEIREYQYNAEKHALDFDGLMSLSAQAGDVVLLHGC 180
QY 181 CHNPTGIDPTLEQWTLAQLSVEKGWLPDFDAYQGFARGLEEDAEGLRAFAMHKLIV 240
DB 181 CHNPTGIDPTAEQWQKLADLSAANGWLPVDFDAYQGFARSLEDEDAEGLRIPTKHNELIV 240
QY 241 ASSYSKNFGLNERNVAGCTLVADSETVDRAFSPQKAAIRANYSNPPAHGASVATILSN 300
DB 241 ASSYSKNFGLNERNVAGCTIVATSDTAKAFSAQKISVRTNYSNPPAHGASVVTILSN 300
QY 301 DALRAIWQEQLTDMRQRTORMRQVNTLOEKGANRDFSLIKONGMFSFGLTKQVLR 360
DB 301 DEBFAEWQELATMRERIRRMQLFVNTLOEKGAKQDFSLISQNGMFSFGLTKQVDR 360
QY 361 LREEFGVAVASGRVNVAGMTDNNAPLCEAIVAVL 396
DB 361 LREEFGIYAVSSGRINVAGLTLENVPLCEAIVAVL 396

RESULT 9
ADN17578
ID ADN17578 standard; protein; 322 AA.
XX
AC ADN17578;
XX

DT 02-DEC-2004 (first entry)
XX Bacterial polypeptide #231.
DE
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW call cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
XX 18-DEC-2003.
PD
XX 20-FEB-2003; 2003US-00369493.
PF
XX 21-FEB-2002; 2002US-0360039P.
PR
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 231; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 322 AA;

Query Match 64.4%; Score 1318; DB 8; Length 322;
Best Local Similarity 76.4%; Pred. No. 5e-124;
Matches 246; Conservative 32; Mismatches 44; Indels 0; Gaps 0;
QY 75 GRCTQELLFGKGSALINDKARTATQGTGTCALRVADFLAKNTSVKRVWISNPSWPHK 134
DB 1 GRVIELLFGSTSAIVTEKARTVQSPGTCALRTAADFIAKQNAKRVWISNPTWPHK 60

QY 135 SVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAQGVVLFPHGCCNPTGIDPTLEQW 194
 Db 61 GVFSAGLEIREYNYDAEKHALNFEGLASLSEAQAQGVVLLHGCHNPTGIDPTPEQW 120
 QY 195 OTLAQLSVEKWLPLDFAYQGFARGLEDEAGLRAPAAHKEELIVASSYKNGFLYNER 254
 Db 121 OKLADLSAANGWLPLDFAYQGFARGLEDEAGLRAPAAHKEELIVASSYKNGFLYNER 180
 QY 255 VGACTLVAADSETVDRAFSQMAAIIANYSNPPAHGASVVATILSNDALRAIWEQELTDM 314
 Db 181 VGACTLVAADSDTAEKAFSQAKFIVNTNYSNPPAHGASVVATILSNEELKAEWQELATM 240
 QY 315 RRIQRMQLFVNTLQEKGNRDFPIIKONGMFSFGLTKBOVLRLREFFGVYAVASGR 374
 Db 241 RRIQRMQLFVNTLQEKAKQDFSFICQNGMFSFGLTKBOVLRLREFFGVYAVASGR 300
 QY 375 VNVAGTTPDNMAPLCEAIVAVL 396
 Db 301 INVAGLTLENMPLCEAIVAVL 322

RESULT 10
 ABP80003
 ID ABP80003 standard; protein; 423 AA.
 XX
 AC ABP80003;
 DT 07-MAR-2003 (first entry)
 DE N. gonorrhoeae amino acid sequence SEQ ID 6536.
 KW Antibacterial; infection; vaccine; gene therapy.
 XX
 OS Neisseria gonorrhoeae.
 XX
 PN WO200279243-A2.
 PD 10-OCT-2002.
 PF 12-FEB-2002; 2002WO-1B002069.
 PR 12-FEB-2001; 2001GB-00003424.
 PA (CHIR-) CHIRON SPA.
 PI Fontana MR, Pizza M, Masignani V, Monaci E;
 XX WPI; 2003-058415/05.
 DR N-PSDB; ABZ40973.
 XX
 PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 XX medicament for treating or preventing N. gonorrhoeae infection.
 PS Disclosure; Page 663; 815pp; English.
 CC The present invention relates to proteins from Neisseria gonorrhoeae.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
 CC molecules of the invention
 XX
 SQ Sequence 423 AA;

Query Match 64.4%; Score 1317; DB 6; Length 423;
 Best Local Similarity 61.9%; Pred. No. 9.8e-124;
 Matches 244; Conservative 62; Mismatches 88; Indels 0; Gaps 0;
 QY 2 FENITAAPADPILGLAIDFRADERPGKINLGIYGVYKBTGKTPVLTSVKKAPQYLLNET 61
 Db 30 FKHEAAPADPILGLGEAFKAEITRPEKVNGLIGVYKDGATPIVKAKEKRLLEST 89

QY 62 TKNYLIDGIPFGRCTQELLFGKGSALINDKRARTAOPTGGTALRVAADFIAKNTSVK 121
 Db 90 TKNYLIDGVADYNEQTQILLFGKDEHIIASRAKTAQSLGGTALRIABFAKRLNAQ 149
 QY 122 RVVNSNPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLNEAQAQGVVLFPHGCC 181
 Db 150 TIWISNPTWPNHNAIAKAVGIQKPYRYDAAKHCLDMDGMIEDLNQAQKGDVILLHGCC 209
 QY 182 HNPTGIDPTLEQWOTLAQLSVEKWLPLDFAYQGFARGLEDEAGLRAPAAHKEELIVA 241
 Db 210 HNPTGIDPTPEQWETLAKLSAEKWLPLDFAYQGFNGLEEDDAYGLRVFLKHNTLELLA 269
 QY 242 SSKSNFGLYNERVGAETLVAADSETVDRAFSQMAAIIANYSNPPAHGASVVATILSND 301
 Db 270 SSKSNFGLYNERVGAETLVAEDEETAARASQIKTIITLTSNPASHGANTIALVLKND 329
 QY 302 ALRAIWEQELTDMRRIQRMQLFVNTLQEKGNRDFPIIKONGMFSFGLTKBOVLRL 361
 Db 330 DLKAQWIAELEDGRKAMRKQFVLLKAKGSTQDFDPIEQNGMFSFGLTKBOVDRL 389
 QY 362 REEFGVYAVASGRVNVAGTTPDNMAPLCEAIVAV 395
 Db 390 KNEFAIYAVRSGRINVAGITDDNIDYLCESIYKV 423

RESULT 11
 AAU73004
 ID AAU73004 standard; protein; 397 AA.
 XX
 AC AAU73004;
 DT 12-MAR-2002 (first entry)
 DE Neisseria meningitidis virulence protein #94.
 KW Meningitis; virulence; gene; antibacterial; vaccine; veterinary;
 KW infection; Gram-negative bacteria; antimicrobial.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO200185772-A2.
 PD 15-NOV-2001.
 PF 08-MAY-2001; 2001WO-GB002003.
 PR 08-MAY-2000; 2000GB-00011108.
 XX (MICR-) MICROSCIENCE LTD.
 PI Tang C;
 XX
 DR WPI; 2002-066593/09.
 DR N-PSDB; AAS97289.
 XX
 PT New peptide encoded by operon including virulence genes of Neisseria
 PT meningitidis, useful as vaccine component for treating or preventing
 PT meningitis and for identifying antimicrobial drug.
 XX
 PS Claim 4; Page 381-383; 423pp; English.
 CC The invention relates to a peptide (I) encoded by an operon (II) of
 CC Neisseria meningitidis including virulence genes, or a related molecule
 CC having a 40% sequence similarity at the peptide or nucleotide level in a
 CC Gram-negative bacterium, or its functional fragment, for therapeutic or
 CC diagnostic use. (I) and (II) are useful in the manufacture of a
 CC medicament for treating or preventing a condition (e.g., meningitis)
 CC associated with infection by Neisseria or Gram-negative bacteria. The
 CC product is useful for veterinary treatment and in a screening assay for
 CC the identification of an antimicrobial drug. The vaccines have
 CC prophylactic applications. AAU72911-AAU73014 represent N. meningitidis
 CC virulence proteins of the invention

```

XX Sequence 397 AA;
SQ
  Query Match      64.1%; Score 1311; DB 5; Length 397;
  Best Local Similarity 61.8%; Pred. No. 3.6e-123;
  Matches 244; Conservative 62; Mismatches 89; Indels 0; Gaps 0;

QY 2 FENITAPADPILGLADLFRADERPGKINLIGVYKDTGKTPVLTTSVKAEQVLLNET 61
DB 3 FKHIAPADPILGLGEAFKAETREPKVNLGIGVYKDSAGATPIVKAKEAKLLESET 62

QY 62 TKNYLIGIDGPEFGRCCTQELLFGKGSALINDKRAARTAQTPGCTGALRVAADFLAKNTSVK 121
DB 63 TKNYLTIDGVADYNAQTQILLFGKHDEIIASRRAKTAQSLGCTGALRIAFAFAKQLNAQ 122

QY 122 RVVSNPSPWNHKSFNVSAGLEVREYAYDAENHTLDFDALINSINEAQGDVVLFGHCC 181
DB 123 TIWISNPTWPNHNAIAKAVGIQDPYRYDDAKHGLDWDGMIEDLSQAQKGDIVLLHGCC 182

QY 182 HNPTGIDPTLEQWTLAQLSVEKWLPLDFAYQGFARGLEDAEGLRAFAAMHKLIVA 241
DB 183 HNPTGIDPTPEQWETLAKLSAEKWLPLDFAYQGFNGLEBEDAYGLRVFLKHNTPELLIA 242

QY 242 SSYSKNFGLYNERVAGCTLVAAADSETVDRAFSPOMKAAIRANYSNPPAHGASVATILSND 301
DB 243 SSYSKNFGMYNERVAGFTLVAAEDEATAARAHQSVKTIIRTLYSNPASHGANTIALVLKND 302

QY 302 ALRAIWEQELTDMRQRIQRMQLFVNTLQEKGANRDFSFIKQNGMFSFSGLTKEQVLR 361
DB 303 DLKAQWIAELDEMGRKAMQKQFVELLKAKGATQDFDFFIEQNGMFSFSGLTPEQVDR 362

QY 362 REEFGVYVAGSRVNVAGTDPDNPALCEAIVAVL 396
DB 363 KNEFAIYVRSGRINVAGITDDNIDYLCESIVKVL 397

RESULT 12
ADP08222
ID ADP08222 standard; protein; 397 AA.
XX
AC ADP08222;
XX
XX ADP08222;
XX
DT 26-AUG-2004 (first entry)
XX
DE Neisseria meningitidis MC58 OMV-related membrane protein - SEQ ID 55.
KW outer-membrane vesicle; antibacterial; antiinflammatory;
KW meningococcal protein trafficking; localisation; infection; vaccine;
KW Gene therapy.
XX
OS Neisseria meningitidis MC58.
XX
PN WO2004046177-A2.
XX
PD 03-JUN-2004.
XX
PF 17-NOV-2003; 2003WO-IB006281.
XX
PR 15-NOV-2002; 2002GB-00026734.
XX
PR 27-MAR-2003; 2003GB-00007131.
XX
XX (CHIR ) CHIRON SRL.
XX
PI Norais N, Grandi G;
XX
XX WPI; 2004-420615/39.
XX
XX New compositions having outer-membrane vesicles and proteins from
PT Neisseria meningitidis, useful in the field of meningococcal
PT biochemistry, in particular for preventing and/or treating meningococcal
PT infections.
XX
PS Claim 9; SEQ ID NO 55; 79pp; English.

```

```

XX The invention relates to a novel composition comprising outer-membrane
CC vesicles (OMV) prepared from a first strain of Neisseria meningitidis and
CC 1 or more proteins which are present in OMVs prepared from a second
CC strain of N. meningitidis, but which are not present in OMVs prepared
CC from the first strain. The composition of the invention demonstrates
CC antibacterial and antiinflammatory activities and may be useful in the
CC field of meningococcal biochemistry, in particular the trafficking and
CC localisation of meningococcal proteins, as well as in the prevention or
CC treatment of meningococcal infections, possibly via the production of a
CC vaccine or gene therapy. The current sequence is that of a Neisseria
CC meningitidis MC58 outer-membrane vesicle (OMV)-related membrane protein
CC of the invention.
XX
SQ Sequence 397 AA;
  Query Match      64.1%; Score 1310; DB 8; Length 397;
  Best Local Similarity 61.8%; Pred. No. 4.5e-123;
  Matches 244; Conservative 64; Mismatches 87; Indels 0; Gaps 0;

QY 2 FENITAPADPILGLADLFRADERPGKINLIGVYKDTGKTPVLTTSVKAEQVLLNET 61
DB 3 FKHIAPADPILGLGEAFKAETREPKVNLGIGVYKDSAGATPIVKAKEAKLLESET 62

QY 62 TKNYLIGIDGPEFGRCCTQELLFGKGSALINDKRAARTAQTPGCTGALRVAADFLAKNTSVK 121
DB 63 TKNYLTIDGVADYNAQTQILLFGKHDEIIASRRAKTAQSLGCTGALRIAFAFAKQLNAQ 122

QY 122 RVVSNPSPWNHKSFNVSAGLEVREYAYDAENHTLDFDALINSINEAQGDVVLFGHCC 181
DB 123 TIWISNPTWPNHNAIAKAVGIQDPYRYDDAKHGLDWDGMIEDLSQAQKGDIVLLHGCC 182

QY 182 HNPTGIDPTLEQWTLAQLSVEKWLPLDFAYQGFARGLEDAEGLRAFAAMHKLIVA 241
DB 183 HNPTGIDPTPEQWETLAKLSAEKWLPLDFAYQGFNGLEBEDAYGLRVFLKHNTPELLIA 242

QY 242 SSYSKNFGLYNERVAGCTLVAAADSETVDRAFSPOMKAAIRANYSNPPAHGASVATILSND 301
DB 243 SSYSKNFGMYNERVAGFTLVAAEDEATAARAHQSVKTIIRTLYSNPASHGANTIALVLKND 302

QY 302 ALRAIWEQELTDMRQRIQRMQLFVNTLQEKGANRDFSFIKQNGMFSFSGLTKEQVLR 361
DB 303 DLKAQWIAELDEMGRKAMQKQFVELLKAKGASQNFDFIIRKQNGMFSFSGLTPEQVDR 362

QY 362 REEFGVYVAGSRVNVAGTDPDNPALCEAIVAVL 396
DB 363 KNEFAIYVRSGRINVAGITDDNIDYLCESIVKVL 397

RESULT 13
ADS22791
ID ADS22791 standard; protein; 394 AA.
XX
AC ADS22791;
XX
XX ADS22791;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #11824.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
XX Bacteria.
XX
OS US2003233675-A1.
XX
XX 18-DEC-2003.
XX
PD
XX

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2006, 02:34:21 ; Search time 1391 Seconds
(without alignments)

5706.433 Million cell updates/sec

Title: US-10-673-786A-1

Perfect score: 1191

Sequence: 1 atgttgagaacattaccgc.....cgattggcagtgctgtaa 1191

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- N_Geneseq_21.*
- 1: Geneseqn1980s.*
 - 2: Geneseqn1990s.*
 - 3: Geneseqn2000s.*
 - 4: Geneseqn2001as.*
 - 5: Geneseqn2001bs.*
 - 6: Geneseqn2002as.*
 - 7: Geneseqn2002bs.*
 - 8: Geneseqn2003as.*
 - 9: Geneseqn2003bs.*
 - 10: Geneseqn2003cs.*
 - 11: Geneseqn2003ds.*
 - 12: Geneseqn2004as.*
 - 13: Geneseqn2004bs.*
 - 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1191	100.0	1191	9 ADB83284	Ad83284 Escherich
2	1191	100.0	1191	13 Adb46042	Ad846042 Bacterial
3	1191	100.0	1191	14 ADW95403	Adw95403 Nucleotid
4	1191	100.0	1331	2 AAV40259	Aav40259 Escherich
5	1191	100.0	3659	1 AAN71109	Aan71109 pheA arof
6	1191	100.0	14759	4 AAS46273	Aas46273 DNA encod
7	1187.8	99.7	1293	1 AAN71108	Aan71108 Optimized
8	839	70.4	1209	11 ACH99107	Ach99107 Klebsiell
9	691.8	58.1	1191	10 ACF68649	Acf68649 Phototrab
10	691.8	58.1	110000	10 ACF67367_14	Continuation (15 o
11	691.8	58.1	249878	10 ACF65381	Acf65381 Phototrab
12	655.8	55.1	1188	13 ADT46374	Adt46374 Bacterial
13	510	42.8	966	13 ADS45488	Ads45488 Bacterial
14	496.2	41.7	2133	5 AAS88191	Aas88191 DNA encod
15	496.2	41.7	3222	5 AAS90082	Aas90082 DNA encod
16	493.4	41.4	7977	13 ADT05500	Adt05500 Haemophil
17	493.4	41.4	349980	13 ADT05649	Adt05649 Haemophil
18	471.8	39.6	1191	6 AAS97289	Aas97289 Neisseria
19	466.8	39.2	17381	3 AAA81493	Aaa81493 N. mening

C	20	466.8	39.2	110000	3	AAA81490_05	Continuation (6 of
C	21	466.8	39.2	349980	3	AAF21607_	Aaf21607 Neisseria
C	22	461.4	38.7	1269	10	ABZ40973	Abz40973 N. gonorr
C	23	458.2	38.5	110000	2	AAT42063_16	Continuation (17 o
C	24	387.6	32.5	579	11	ACH99159	Ach99159 Klebsiell
C	25	372	31.2	373	4	AAS41443	Aas41443 cDNA enco
C	26	363	30.5	2220	5	AAS81804	Aas81804 DNA encod
C	27	348.4	29.3	702	10	ADF03520	Adf03520 Bacterial
C	28	331.8	27.9	1185	13	ADS59837	Ads59837 Bacterial
C	29	327.8	27.5	1191	13	ADS56437	Ads56437 Bacterial
C	30	320.8	26.9	1188	13	ADS55510	Ads55510 Bacterial
C	31	320.8	26.9	1194	13	ADS49995	Ads49995 Bacterial
C	32	316.4	26.6	1509	11	ABD14364	Abd14364 Pseudomon
C	33	316.4	26.6	1632	11	ABD13903	Abd13903 Pseudomon
C	34	310.6	26.1	1194	13	ADS62038	Ads62038 Bacterial
C	35	306.6	25.7	24417	2	AAT97221	Aat97221 Pseudomon
C	36	305	25.6	1242	11	ABD13824	Abd13824 Pseudomon
C	37	305	25.6	1329	11	ABD13558	Abd13558 Pseudomon
C	38	297.4	25.0	1093	5	AAS77350	Aas77350 DNA encod
C	39	289	24.3	1338	11	ACH95506	Ach95506 Klebsiell
C	40	288.8	24.2	1188	13	ADS56672	Ads56672 Bacterial
C	41	287.6	24.1	1185	13	ADS61736	Ads61736 Bacterial
C	42	285.4	24.0	1191	10	ABZ38235	Abz38235 N. gonorr
C	43	285.2	23.9	1191	13	ADS63841	Ads63841 Bacterial
C	44	285.2	23.9	1191	13	ADS64221	Ads64221 Bacterial
C	45	285.2	23.9	1194	13	ADS63473	Ads63473 Bacterial

ALIGNMENTS

RESULT 1	
ID	ADB83284
ID	ADB83284 standard; DNA; 1191 BP.
XX	
AC	ADB83284;
XX	
DT	04-DEC-2003 (first entry)
XX	
DE	Escherichia coli aspartate aminotransferase gene.
XX	
KW	ds; gene; aspartate aminotransferase; threonine; fermentation.
XX	
OS	Escherichia coli.
XX	
PH	Key
FT	CDS
FT	1..1191
FT	/tag= a
FT	/product= "aspartate aminotransferase"
XX	
PN	WO2003072786-A1.
XX	
PD	04-SEP-2003.
XX	
XX	25-FEB-2003; 2003WO-JP002067.
PF	
XX	
PR	27-FEB-2002; 2002RU-00104983.
XX	
PA	(AJIN) AJINOMOTO CO INC.
XX	
PI	Akhverdian VZ, Savrasova EA, Kaplan AM, Lobanov AO, Kozlov YI;
XX	
DR	WPI; 2003-721777/68.
DR	P-FSD; ADB83285.
XX	
PT	Industrial production of L-threonine by fermentation using Escherichia
PT	modified to enhance aspartate aminotransferase activity, with improved
PT	productivity.
XX	
PS	Claim 7; Page 18-20; 26pp; Japanese.
XX	
CC	The invention relates to a bacterium belonging to the genus Escherichia
CC	which is modified to enhance aspartate aminotransferase activity for the

CC production of L-threonine. The bacterium is used for the industrial
CC production of L-threonine by fermentation. This sequence corresponds to
CC the E. coli aspartate aminotransferase gene.
XX
SQ Sequence 1191 BP; 291 A; 305 C; 317 G; 278 T; 0 U; 0 Other;
Query Match 100.0%; Score 1191; DB 9; Length 1191;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTTGGAGAACTTACCGCGCTCTCGCGACCCGATTCCTGGCGCTGGCCGATCTGTTT 60
DB 1 ATGTTTGGAGAACTTACCGCGCTCTCGCGACCCGATTCCTGGCGCTGGCCGATCTGTTT 60
QY 61 CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTCTCTATAAAGATGAGCG 120
DB 61 CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTCTCTATAAAGATGAGCG 120
QY 121 GGCANAACCCGCTACTGACCGCTGAAAGAGCTGACAGTATCTGCTCGAANAATGAA 180
DB 121 GGCANAACCCGCTACTGACCGCTGAAAGAGCTGAAAGAGCTGAAAGATGAA 180
QY 181 ACCACCAAAATTTACCTCGGCATTTGACGGCATCCCTGGAATTTGGTGGCTGCACTCAGGAA 240
DB 181 ACCACCAAAATTTACCTCGGCATTTGACGGCATCCCTGGAATTTGGTGGCTGCACTCAGGAA 240
QY 241 CTGCTGTTGGTAAAGTAGCGCTGATCAATGACAAACGTCCTCGCACGGCACAGACT 300
DB 241 CTGCTGTTGGTAAAGTAGCGCTGATCAATGACAAACGTCCTCGCACGGCACAGACT 300
QY 301 CCGGGGGCACTGGCGCACTACCGTGGCTGCGGATTTCTGCAAAAATACACGGTT 360
DB 301 CCGGGGGCACTGGCGCACTACCGTGGCTGCGGATTTCTGCAAAAATACACGGTT 360
QY 361 AACCGTGTGGTGAGCAACCAAGCTGGCGCAACCAATGAAGCGTCTTTAACTCTGCA 420
DB 361 AACCGTGTGGTGAGCAACCAAGCTGGCGCAACCAATGAAGCGTCTTTAACTCTGCA 420
QY 421 GGTCTGGAAGTTCGTGAATACGTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
DB 421 GGTCTGGAAGTTCGTGAATACGTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
QY 481 GCATGATTAACGCTGATGAAGCTCAGGCTGGCGAGTGTGTTCCATGGCTGC 540
DB 481 GCATGATTAACGCTGATGAAGCTCAGGCTGGCGAGTGTGTTCCATGGCTGC 540
QY 541 TGCATTAACCAACCGGTATGACCCCTACGCTGGAACAATGCAACACATGGCACAACTC 600
DB 541 TGCATTAACCAACCGGTATGACCCCTACGCTGGAACAATGCAACACATGGCACAACTC 600
QY 601 TCCGTTTGAGAAAGCTGGTTACCGCTGTTTGAATTCGCTTACCAGGGTTTGGCCGTGGT 660
DB 601 TCCGTTTGAGAAAGCTGGTTACCGCTGTTTGAATTCGCTTACCAGGGTTTGGCCGTGGT 660
QY 661 CTGGAAGAACTGCTGAAGACTGCGCGCTTTGCGGCTATGATGAAGAGCTGATGTT 720
DB 661 CTGGAAGAACTGCTGAAGACTGCGCGCTTTGCGGCTATGATGAAGAGCTGATGTT 720
QY 721 GCCAGTTCCTACTCTAAAACTTTGGCTGTACACGAGCGTGTGGCGCTTGTACTCTG 780
DB 721 GCCAGTTCCTACTCTAAAACTTTGGCTGTACACGAGCGTGTGGCGCTTGTACTCTG 780
QY 781 GTTGCTGCCGACAGTGAACCGTTGATTCGCGCATTCAGCCAAATGAAGCGCGATTCGC 840
DB 781 GTTGCTGCCGACAGTGAACCGTTGATTCGCGCATTCAGCCAAATGAAGCGCGATTCGC 840
QY 841 GCTAACTACTCTAACCCACAGACACGCGGCTTCTGTTGCGCCACCATCTTGAGCAAC 900
DB 841 GCTAACTACTCTAACCCACAGACACGCGGCTTCTGTTGTTGCCACCATCTTGAGCAAC 900
QY 901 GATGCGTTACGTGGATTTGGGAACAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT 960
DB 901 GATGCGTTACGTGGATTTGGGAACAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT 960

QY 961 ATGCGTTCAGTTCTTCTGTAATACGCTGCAGGAAAAAGCGCAAAACCGGACTTCAGCTTT 1020
DB 961 ATGCGTTCAGTTCTTCTGTAATACGCTGCAGGAAAAAGCGCAAAACCGGACTTCAGCTTT 1020
QY 1021 ATCATCAAAACAGACGCGCATGTTCTCTTCAGTGGCCTGACAAAAGAAACAAGTGTGCGT 1080
DB 1021 ATCATCAAAACAGACGCGCATGTTCTCTTCAGTGGCCTGACAAAAGAAACAAGTGTGCGT 1080
QY 1081 CTGCGCGAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGGTAATGTGGCCGGATG 1140
DB 1081 CTGCGCGAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGGTAATGTGGCCGGATG 1140
QY 1141 ACACAGATAAATCATGCTCCGCTGTCGAAGCGATTGTGGCAGTGTCTGTA 1191
DB 1141 ACACAGATAAATCATGCTCCGCTGTCGAAGCGATTGTGGCAGTGTCTGTA 1191
RESULT 2
ADS46042 standard; cDNA; 1191 BP.
XX AC ADS46042;
XX 02-DEC-2004 (first entry)
XX Bacterial polynucleotide #785.
DE Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX Bacteria.
XX US2003233675-A1.
XX 18-DEC-2003.
XX 20-FEB-2003; 2003US-00369493.
XX 21-FEB-2002; 2002US-0360039P.
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
DR New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX Claim 1; SEQ ID NO 24472; 122pp; English.
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ

Query Match 100.0%; Score 1191; DB 13; Length 1191;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGTTTGAGAACATTACCGCGCTCTCGCGACCCGATTCCTGGCCCTGGCCGATCTGTTT 60
 DB 1 ATGTTTGAGAACATTACCGCGCTCTCGCGACCCGATTCCTGGCCCTGGCCGATCTGTTT 60
 QY 61 CTGTCGCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGTTCTATAAAGATGAGACG 120
 DB 61 CTGTCGCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGTTCTATAAAGATGAGACG 120
 QY 121 GGCAAAACCCCGTACTACGACGGTGGAAGAGGCTGAACAGTATCTGCTCGAAATGAA 180
 DB 121 GGCAAAACCCCGTACTACGACGGTGGAAGAGGCTGAACAGTATCTGCTCGAAATGAA 180
 QY 181 ACCACCAAAATTAACCTCGGATTTGAGCGCATCCCTGAAATTTGCTGCTGCACCTCAGGAA 240
 DB 181 ACCACCAAAATTAACCTCGGATTTGAGCGCATCCCTGAAATTTGCTGCTGCACCTCAGGAA 240
 QY 241 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGCTGCTCGACGGCACAGACT 300
 DB 241 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGCTGCTCGACGGCACAGACT 300
 QY 301 CCGGGGGGCACTGGCGCACTACCGTGCGCTGCGGATTTCTGGCAAAAATACACAGGTT 360
 DB 301 CCGGGGGGCACTGGCGCACTACCGTGCGCTGCGGATTTCTGGCAAAAATACACAGGTT 360
 QY 361 AAGCGTGTTGGTGAGCAACCAAGCTGGCCCAACCAATGAAGCGCTTTAACTCTGCA 420
 DB 361 AAGCGTGTTGGTGAGCAACCAAGCTGGCCCAACCAATGAAGCGCTTTAACTCTGCA 420
 QY 421 GGTCTGGAAGTTCTGTAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
 DB 421 GGTCTGGAAGTTCTGTAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
 QY 481 GCATGATTAACAGCTGAATGAAGCTAGGCTGGCGACGATGCTGTTTCCATGCTGTC 540
 DB 481 GCATGATTAACAGCTGAATGAAGCTAGGCTGGCGACGATGCTGTTTCCATGCTGTC 540
 QY 541 TGCCATAACCCACCGGTATCGACCTACGCTGGAACCAATGGCAACACTGGCACAACTC 600
 DB 541 TGCCATAACCCACCGGTATCGACCTACGCTGGAACCAATGGCAACACTGGCACAACTC 600
 QY 601 TCGTTTGAGAAAGGCTGGTTACCGCTGTTTGAATCTCGCTTACCAGGGTTTTCGCCGTGGT 660
 DB 601 TCGTTTGAGAAAGGCTGGTTACCGCTGTTTGAATCTCGCTTACCAGGGTTTTCGCCGTGGT 660
 QY 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCCGCGCTATGATGAAGAGCTGATGTT 720
 DB 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCCGCGCTATGATGAAGAGCTGATGTT 720
 QY 721 GCGAGTTCTCTCTAAACCTTTGGCTGTACCAACGAGCGTTTGGCGCTTGATCTCTG 780
 DB 721 GCGAGTTCTCTCTAAACCTTTGGCTGTACCAACGAGCGTTTGGCGCTTGATCTCTG 780
 QY 781 GTTGCTGCCGACAGTGAACCCGTTTGTATCGCGCATTCAGCCAAATGAAGCGCGATTTCG 840

DB 781 GTTGCTGCCGACAGTGAACCCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTTCG 840
 QY 841 GCTAACTACTCTAACCCACGACGCGGCTTCTGTTGTCACCATCTCTGAGCAAC 900
 DB 841 GCTAACTACTCTAACCCACGACGCGGCTTCTGTTGTCACCATCTCTGAGCAAC 900
 QY 901 GATCGCTTACGTCGATTTGGGAAACAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT 960
 DB 901 GATCGCTTACGTCGATTTGGGAAACAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT 960
 QY 961 ATGCGTCAGTTTGTGCTCAATACGTCGACGAAAAAGCGCAACACCGCACTTCAGCTTT 1020
 DB 961 ATGCGTCAGTTTGTGCTCAATACGTCGACGAAAAAGCGCAACACCGCACTTCAGCTTT 1020
 QY 1021 ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCTGACAAAGACAAAGTCTGCGT 1080
 DB 1021 ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCTGACAAAGACAAAGTCTGCGT 1080
 QY 1081 CTGCGGAAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTGGCCGGATG 1140
 DB 1081 CTGCGGAAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTGGCCGGATG 1140
 QY 1141 ACACGAGATAACATGCTCGCTGTCGAAAGCGATTGTGCGCAGTGTCTGTAA 1191
 DB 1141 ACACGAGATAACATGCTCGCTGTCGAAAGCGATTGTGCGCAGTGTCTGTAA 1191
 RESULT 3
 ADM95403
 ID ADM95403 standard; DNA; 1191 BP.
 XX
 AC ADM95403;
 XX
 DT 07-APR-2005 (first entry)
 XX
 DE Nucleotide sequence of an Escherichia coli polypeptide.
 XX
 KW transgenic plant; glutamic acid dehydrogenase; GDH; ECASPC;
 KW 2-oxo glutaric acid; agriculture; gene; ds.
 XX
 OS Escherichia coli.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1191
 FT /*tag= a
 XX
 PN WO2005006847-A1.
 XX
 PD 27-JAN-2005.
 XX
 PF 15-JUL-2004; 2004WO-JP010451.
 XX
 PR 17-JUL-2003; 2003JP-00198559.
 XX
 PA (AJIN) AJINOMOTO CO INC.
 XX
 PI Kisaka H, Miwa T, Akiyama A;
 XX
 DR WPI; 2005-132242/14.
 DR P-PSDB; ADM95404.
 XX
 PT Producing plant having improved growth and yield under cultivation
 PT conditions with decrease in nitrogen, involves expressing transduced
 PT glutamic acid dehydrogenase gene in plant so as to increase their 2-oxo
 PT glutaric acid contents.
 XX
 PS Disclosure; SEQ ID NO 19; 94pp; Japanese.
 XX
 CC The specification describes a method of producing a plant having improved
 CC growth and yield under cultivation conditions with a decrease in
 CC nitrogen. The method involves transducing glutamic acid dehydrogenase
 CC (GDH) or ECASPC genes into a plant and expressing the gene so as to

CC increase 2-oxo glutaric acid content of the plant, or applying proline to
CC the foliage of the plant so as to increase the 2-oxo glutaric acid
CC content of the plant. The method is useful for producing a plant having
CC improved growth and yield under cultivation conditions with a decrease in
CC nitrogen. The present sequence encodes a polypeptide, used in the course
CC of the invention.
XX
SQ Sequence 1191 BP; 291 A; 305 C; 317 G; 278 T; 0 U; 0 Other;

Query Match 100.0%; Score 1191; DB 14; Length 1191;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTGAGAACATTACCGCGCTCTCGCGACCCCGAATCTGGGCCCTGGCCGATCTGTTT 60
DB 1 ATGTTTGAGAACATTACCGCGCTCTCGCGACCCCGAATCTGGGCCCTGGCCGATCTGTTT 60
QY 61 CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGCTCTATAAGATGAGCG 120
DB 61 CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGCTCTATAAGATGAGCG 120
QY 121 GGCACAAACCCGGTACTGACCGCGTGAAGAGGCTGAACAGATATCTGCTCGAAAAATGAA 180
DB 121 GGCACAAACCCGGTACTGACCGCGTGAAGAGGCTGAACAGATATCTGCTCGAAAAATGAA 180
QY 181 ACCACCAAAATTAACCTCGGATTTGACGGGATCCCTGAATTTGGTGGCTGCATCAGGAA 240
DB 181 ACCACCAAAATTAACCTCGGATTTGACGGGATCCCTGAATTTGGTGGCTGCATCAGGAA 240
QY 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCACGGCAGACT 300
DB 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCACGGCAGACT 300
QY 301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGATTTCTTGGCAAAAATACCAAGCGTT 360
DB 301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGATTTCTTGGCAAAAATACCAAGCGTT 360
QY 361 AAGCGTGTGGTGAGCAACCCAGCTGCGCAACATTAAGAGCGTCTTTAACTCTGCA 420
DB 361 AAGCGTGTGGTGAGCAACCCAGCTGCGCAACATTAAGAGCGTCTTTAACTCTGCA 420
QY 421 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACAACCTTTCGATTCGAT 480
DB 421 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACAACCTTTCGATTCGAT 480
QY 481 GCACTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTTCCATGGCTGC 540
DB 481 GCACTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTTCCATGGCTGC 540
QY 541 TGCCATAACCCAAACCGGTATCGACCCCTACGCTGGAACAATGGCAAAACACATGGCACAATC 600
DB 541 TGCCATAACCCAAACCGGTATCGACCCCTACGCTGGAACAATGGCAAAACACATGGCACAATC 600
QY 601 TCCGTTTGAGAAAGGCTGTTTACCGCTGTTTGAATTCGCTTACCAGGTTTTCGCCGTGT 660
DB 601 TCCGTTTGAGAAAGGCTGTTTACCGCTGTTTGAATTCGCTTACCAGGTTTTCGCCGTGT 660
QY 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGCTATGCAATGAAGAGCTGATTGTT 720
DB 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGCTATGCAATGAAGAGCTGATTGTT 720
QY 721 GCCAGTTCCTACTCTAAAAAATTTGGCCTGTACACAGGCGTGTGGCGCTTTGACTCTG 780
DB 721 GCCAGTTCCTACTCTAAAAAATTTGGCCTGTACACAGGCGTGTGGCGCTTTGACTCTG 780
QY 781 GTTGCTGCCGACAGTGAACCGTTGATCGCGATTTCAGCCAAATGAAGCGGCGATTTCG 840
DB 781 GTTGCTGCCGACAGTGAACCGTTGATCGCGATTTCAGCCAAATGAAGCGGCGATTTCG 840
QY 841 GCTAACTACTCTAACCCACCAGCACAGCGGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 900
DB 841 GCTAACTACTCTAACCCACCAGCACAGCGGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 900

QY 901 GATGCTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCGCAGCGTATTGAGCGT 960
DB 901 GATGCTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCGCAGCGTATTGAGCGT 960
QY 961 ATGCGTCAAGTTGTTTCGTCAATACGCTGAGGAAAAAGCGCAAAACCGGACTTCAGCTTT 1020
DB 961 ATGCGTCAAGTTGTTTCGTCAATACGCTGAGGAAAAAGCGCAAAACCGGACTTCAGCTTT 1020
QY 1021 ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCCTGACAAAAGAACAGTGTGCGCT 1080
DB 1021 ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCCTGACAAAAGAACAGTGTGCGCT 1080
QY 1081 CTGCGCGAAGAGTTTGGCGTATATCGGTTGCTTCTGTCGCTGCTAAATGTGCGCGGATG 1140
DB 1081 CTGCGCGAAGAGTTTGGCGTATATCGGTTGCTTCTGTCGCTGCTAAATGTGCGCGGATG 1140
QY 1141 ACACAGATATAACATGGCTCCGCTGCGAAGCGATTGTGGCAGTGTCTGTA 1191
DB 1141 ACACAGATATAACATGGCTCCGCTGCGAAGCGATTGTGGCAGTGTCTGTA 1191

RESULT 4
AAV40259
ID AAV40259 standard; DNA; 1331 BP.

AC AAV40259;
XX 13-OCT-1998 (first entry)
XX Escherichia coli aspC gene.
XX Brevibacterium lactofermentum; lysC; L-lysine; coryneform bacterium;
KW aspartokinase; feedback inhibition; dihydropicolinate reductase;
KW diaminopimelate decarboxylase; aspartate aminotransferase; ds.
XX Escherichia coli.

Key	Location/Qualifiers
FT CDS	10..1197
FT	/*tag= a
FT	/product= "aspC"
XX	EP854189-A2.
XX	22-JUL-1998.
XX	05-DEC-1997; 97EP-00121443.
XX	05-DEC-1996; 96JP-00325659.
XX	(AJIN) AJINOMOTO CO INC.
XX	Araki M, Sugimoto M, Yoshihara Y, Nakamatsu T;
XX	WPI; 1998-379060/33.
XX	P-PSDB; AAW69553.

Recombinant DNA autonomously replicable in coryneform bacteria - used to produce L-lysine, codes for e.g. aspartokinase, dihydropicolinate reductase and synthase and di:amino-pimelate decarboxylase.
Example 5; Page 37-38; 59pp; English.

The present invention describes a recombinant DNA autonomously replicable in cells of coryneform bacteria (CB), comprising a DNA sequence coding for an aspartokinase (AK) in which feedback inhibition by L-lysine and L-threonine is desensitized, a DNA sequence coding for a dihydropicolinate reductase (DHP), a DNA sequence coding for dihydropicolinate synthase (DHPs), a DNA sequence coding for diaminopimelate decarboxylase (DAMD) and a DNA sequence coding for aspartate aminotransferase (AAT). The present sequence encodes aspC from Escherichia coli. The DNA and related products from the present invention, can be used for improving L-lysine productivity by CB. The L-

CC lysine produced can be used as a fodder additive

XX	Sequence	1331 BP; 330 A; 340 C; 350 G; 311 T; 0 U; 0 Other;	
XX	Query Match	100.0%; Score 1191; DB 2; Length 1331;	
XX	Best Local Similarity	100.0%; Pred. No. 0;	
XX	Matches 1191; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	ATGTTTTCAGAACATTACCGCGCTCCCTGCGGACCCGATTCCTGGGCTTGGCGGATCTGTTT	60
DB	10	ATGTTTTCAGAACATTACCGCGCTCCCTGCGGACCCGATTCCTGGGCTTGGCGGATCTGTTT	69
QY	61	CGTGCCGATGAAGTCCCGGCAAAATTAACCTCGGGATTGGTGCTATAAAGATGAGACG	120
DB	70	CGTGCCGATGAAGTCCCGGCAAAATTAACCTCGGGATTGGTGCTATAAAGATGAGACG	129
QY	121	GGCAAAACCCGGTACTGACGAGGTGAAAAGGCTGAACAGTATCTGCTCGAAAATGAA	180
DB	130	GGCAAAACCCGGTACTGACGAGGTGAAAAGGCTGAACAGTATCTGCTCGAAAATGAA	189
QY	181	ACCACCAAAATTAACCTCGGCAATTCAGCGATCCCTGAATTTGGTCCGCTGCACCTCAGGAA	240
DB	190	ACCACCAAAATTAACCTCGGCAATTCAGCGATCCCTGAATTTGGTCCGCTGCACCTCAGGAA	249
QY	241	CTGCTGTTTGGTAAAGTAGCGCTGATCAATGACAAACGTCGTCGACGCGCACAGACT	300
DB	250	CTGCTGTTTGGTAAAGTAGCGCTGATCAATGACAAACGTCGTCGACGCGCACAGACT	309
QY	301	CCGGGGGCACTGGCGCACTACGCGTGCGTGGCGATTTCTTGGCAAAAAATACAGCGTT	360
DB	310	CCGGGGGCACTGGCGCACTACGCGTGCGTGGCGATTTCTTGGCAAAAAATACAGCGTT	369
QY	361	AAGCGTGTTGGTGACCAACCAAGCTGGCGCAACCAAGCGTCTTAACTCTGCA	420
DB	370	AAGCGTGTTGGTGACCAACCAAGCTGGCGCAACCAAGCGTCTTAACTCTGCA	429
QY	421	GGTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAATCACACTTTGACTTCGAT	480
DB	430	GGTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAATCACACTTTGACTTCGAT	489
QY	481	GCACCTGATTAACAGCTGAATGAAGCTCAGGCTGGCGAAGCTAGTGCTGTTCCATGGCTGC	540
DB	490	GCACCTGATTAACAGCTGAATGAAGCTCAGGCTGGCGAAGCTAGTGCTGTTCCATGGCTGC	549
QY	541	TGCGATAACCAACCGGTATCGACCTACGCTGGCAACAAATGGCAACACTGGCACACTC	600
DB	550	TGCGATAACCAACCGGTATCGACCTACGCTGGCAACAAATGGCAACACTGGCACACTC	609
QY	601	TCGGTTGAGAAAGCGTGTACCGCTGTTTGACTTTGACTTACAGGGTTTTGGCCGCTGGT	660
DB	610	TCGGTTGAGAAAGCGTGTACCGCTGTTTGACTTTGACTTACAGGGTTTTGGCCGCTGGT	669
QY	661	CTGGAAGAAGATGCTGAAGACTCGCGCTTTTCGCGCTATGCAATAAGAGCTGATGTTT	720
DB	670	CTGGAAGAAGATGCTGAAGACTCGCGCTTTTCGCGCTATGCAATAAGAGCTGATGTTT	729
QY	721	GCCAGTTCCTACTCTAATAAATTTGGCTGTACAGAGCGTGTGGCGCTTGTACTCTG	780
DB	730	GCCAGTTCCTACTCTAATAAATTTGGCTGTACAGAGCGTGTGGCGCTTGTACTCTG	789
QY	781	GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGGCGATTCGC	840
DB	790	GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGGCGATTCGC	849
QY	841	GCTAACTACTCTAACCCACGACACACGCGCTTCTGTTGTGCCACCATCTCTGAGCAAC	900
DB	850	GCTAACTACTCTAACCCACGACACACGCGCTTCTGTTGTGCCACCATCTCTGAGCAAC	909
QY	901	GATGCGTTAGTGGGATTTGGGACAAAGACTGACTGATATGCGCCAGCGTATTCAGCGT	960
DB	910	GATGCGTTAGTGGGATTTGGGACAAAGACTGACTGATATGCGCCAGCGTATTCAGCGT	969
QY	961	ATGCGTCAGTTGTTTCGTCAATACGCTGCAAGGAAAAAGGCGCAAAACCGCGACTTCAGCTTT	1020

DB	970	ATGCGTCAGTTGTTTCGTCAATACGCTGCAAGAAAAAGCGCAAAACCGCGACTTCAGCTTT	1029
QY	1021	ATCATCAACAGAACGCGCATGTTCTCCTTCAGTGGCCTGACAAAGAACAGTCTGCGT	1080
DB	1030	ATCATCAACAGAACGCGCATGTTCTCCTTCAGTGGCCTGACAAAGAACAGTCTGCGT	1089
QY	1081	CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTGCGCGGATG	1140
DB	1090	CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTGCGCGGATG	1149
QY	1141	ACACCAAGATTAACATGGCTCCGCTGTGCGAAGCGATTTGCGAGTGTGTTAA	1191
DB	1150	ACACCAAGATTAACATGGCTCCGCTGTGCGAAGCGATTTGCGAGTGTGTTAA	1200
RESULT 5			
XX	AA71109	standard; DNA; 3659 BP.	
XX	AC	AA71109;	
XX	01-JAN-1980	(first entry)	
XX	pheA	aroF aspC operon in plasmid pME219.	
XX	alpha-amylase;	feedback inhibition; amino acid synthesis;	
XX	composite plasmid; ss.		
XX	Bacillus	licheniformis.	
XX	Key	Location/Qualifiers	
XX	CDS	69..1241	
XX	FT	/*tag= a	
XX	CDS	1259..2329	
XX	FT	/*tag= b	
XX	CDS	2344..3534	
XX	FT	/*tag= c	
XX	WO800202-A.		
XX	15-JAN-1987.		
XX	24-JUN-1986;	86WO-US001353.	
XX	24-JUN-1985;	85US-00747732.	
XX	(NUTR-)	NUTRASWEET CO.	
XX	Edwards MR,	Taylor PP, Hunter MG, Fotheringh IG;	
XX	WPI; 1987-021998/03.		
XX	P-PSDB; AAP70752,	AAP71677, AAP71678.	
XX	Composite plasmids	contg. multiple genes in transcriptional units -	
XX	useful for prodn.	of aminoacid(s), esp. L-phenylalanine and l-tyrosine.	
XX	Disclosure;	Page 38; 57pp; English.	
XX	This sequence	may be inserted into a composite plasmid and used for the	
XX	production of amino	acids. See also AA71053-55, AA71107, AA71109,	
XX	AA71111 and AAP70696-97	and AAP70750, AAP70752 and AAP70754	
XX	Sequence	3659 BP; 936 A; 911 C; 947 G; 865 T; 0 U; 0 Other;	
XX	Query Match	100.0%; Score 1191; DB 1; Length 3659;	
XX	Best Local Similarity	100.0%; Pred. No. 0;	
XX	Matches 1191; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	ATGTTTTCAGAACATTACCGCGCTTCCTGCGGACCCGATTCCTGGGCTTGGCGGATCTGTTT	60
DB	2344	ATGTTTTCAGAACATTACCGCGCTTCCTGCGGACCCGATTCCTGGGCTTGGCGGATCTGTTT	2403

QY 61 CGTGCCGATGAACGTCCTCCGCGAATAATTAACTCCGGATTGGTGCTCTATAAAGATGAGACG 120
DB 2404 CGTGCCGATGAACGTCCTCCGCGAATAATTAACTCCGGATTGGTGCTCTATAAAGATGAGACG 2463
QY 121 GGCAAAACCCCGGTACTACACGCGTGAAGAGGCTGAACAGTATCTGCTCGAAAAATGAA 180
DB 2464 GGCAAAACCCCGGTACTACACGCGTGAAGAGGCTGAACAGTATCTGCTCGAAAAATGAA 2523
QY 181 ACCACAAAATAATTACTCTGGCATTTGACGGCATCCCTGAAATTTGGTCTGCTGCACTCAGGAA 240
DB 2524 ACCACAAAATAATTACTCTGGCATTTGACGGCATCCCTGAAATTTGGTCTGCTGCACTCAGGAA 2583
QY 241 CTGCTGTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTGCACGGCACAGACT 300
DB 2584 CTGCTGTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTGCACGGCACAGACT 2643
QY 301 CCGGGGGGCACTGGCGCACTACCGTGGCTGGCGATTCTTCTGGCAAAAAATACCAACGCTT 360
DB 2644 CCGGGGGGCACTGGCGCACTACCGTGGCTGGCGATTCTTCTGGCAAAAAATACCAACGCTT 2703
QY 361 AAGCGTGTGGGTGAGCAACCAAGCTGGCCGAAACATAAGAGCGTCTTTAACTCTGCA 420
DB 2704 AAGCGTGTGGGTGAGCAACCAAGCTGGCCGAAACATAAGAGCGTCTTTAACTCTGCA 2763
QY 421 GGTCTGGAAGTTCTGTAATACGCTTATATGATGCGGAAATCACAACCTCTTGACTTCGAT 480
DB 2764 GGTCTGGAAGTTCTGTAATACGCTTATATGATGCGGAAATCACAACCTCTTGACTTCGAT 2823
QY 481 GCACTGATTAACAGCTGAATGAAGCTCAAGCTGGCGACGCTAGTGTCTGTTCATGCTGC 540
DB 2824 GCACTGATTAACAGCTGAATGAAGCTCAAGCTGGCGACGCTAGTGTCTGTTCATGCTGC 2883
QY 541 TGCATTAACCAACCCGATTCGACCTAGCTCGTGAACAAATGGCAAAACATGGCACAACTC 600
DB 2884 TGCATTAACCAACCCGATTCGACCTAGCTCGTGAACAAATGGCAAAACATGGCACAACTC 2943
QY 601 TCCGTTGAGAAAGGCTGTTACCGCTGTTTGACTTCGCTTACCAGGTTTTCGCCCTGGT 660
DB 2944 TCCGTTGAGAAAGGCTGTTACCGCTGTTTGACTTCGCTTACCAGGTTTTCGCCCTGGT 3003
QY 661 CTGGAAGAGATGCTCAAGAGACTCGCGCTTTCGCGCTATGATAAAGAGCTGATTGTT 720
DB 3004 CTGGAAGAGATGCTCAAGAGACTCGCGCTTTCGCGCTATGATAAAGAGCTGATTGTT 3063
QY 721 GCCAGTTCCTACTCTAAAAATTTGGCTGTACAACGAGCGTTGGCGCTTGTACTCTG 780
DB 3064 GCCAGTTCCTACTCTAAAAATTTGGCTGTACAACGAGCGTTGGCGCTTGTACTCTG 3123
QY 781 GTTGCTGCCGACAGTGAAACCGTTGATCGCGCAATTCAGCCAAATGAAGCGGCGATTCCG 840
DB 3124 GTTGCTGCCGACAGTGAAACCGTTGATCGCGCAATTCAGCCAAATGAAGCGGCGATTCCG 3183
QY 841 GCTAACTACTCTAACCCACGACACACGCGCTTCTGTTGTCACCACTCTCTGAGCAAC 900
DB 3184 GCTAACTACTCTAACCCACGACACACGCGCTTCTGTTGTCACCACTCTCTGAGCAAC 3243
QY 901 GATGCGTTACGTCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960
DB 3244 GATGCGTTACGTCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 3303
QY 961 ATGCGTCAGTTGTTGTCATACGCTGCAGGAAAAGGCGCAAAACCGCACTTCAGCTTT 1020
DB 3304 ATGCGTCAGTTGTTGTCATACGCTGCAGGAAAAGGCGCAAAACCGCACTTCAGCTTT 3363
QY 1021 ATCATCAACAGAACGCGATGTTCTCCTTCAGTGGCTGACAAAAGAACAAAGTGGCTGCGT 1080
DB 3364 ATCATCAACAGAACGCGATGTTCTCCTTCAGTGGCTGACAAAAGAACAAAGTGGCTGCGT 3423
QY 1081 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGCGTAAATGTCGCCGGATG 1140
DB 3424 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGCGTAAATGTCGCCGGATG 3483
QY 1141 ACACAGATAACATGGCTCGCTGTGGGAAGCGATTGTGGCAGTGTCTGTAA 1191

DB 3484 ACACAGATAACATGGCTCGCTCCGCTGTGCGAAGCGATTGTGCGAGTGTCTGTAA 3534
RESULT 6
AAS46273/c
ID AAS46273 standard; DNA; 14759 BP.
XX AAS46273;
XX 18-DEC-2001 (first entry)
XX DNA encoding novel mar regulated protein (NIMR) #42.
XX mar regulated polypeptide; NIMR; microbial infection; antibacterial; ds.
XX Escherichia coli.
XX WO200170776-A2.
XX 27-SEP-2001.
XX 08-MAR-2001; 2001WO-US007478.
XX 10-MAR-2000; 2000US-0188362P.
XX (TUFT) TUFTS COLLEGE.
XX Levy SB, Barbosa TM, Alekshun MN;
XX WPI: 2001-602769/68.
XX P-PSDB; AAU29374.
XX Identifying compounds that modulate a newly identified mar regulated polypeptide activity, useful as antimicrobial compounds, involves contacting the polypeptide with a test compound.
XX Disclosure; Page 477-485; 526pp; English.
XX The invention relates to a method of identifying compounds that modulate a newly identified mar regulated (NIMR) polypeptide activity. The method comprises contacting an NIMR polypeptide with a test compound under interaction conditions, determining the ability of the compound to modulate the activity or expression of the polypeptide, and selecting the modulators. NIMR nucleic acids and polypeptides are used in the treatment of microbial infections, and in screening for modulators of NIMR expression and activity. These modulators can be used to reduce the infectivity of a microbe on a surface, and the virulence of a microbe in a subject suffering from an infection. AAS46232-AAS46278 represent Escherichia coli NIMR coding sequences of the invention
XX Sequence 14759 BP; 3703 A; 3840 C; 3713 G; 3503 T; 0 U; 0 Other;
Query Match 100.0%; Score 1191; DB 4; Length 14759;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTTGTAGAACATTACCGCCGCTCTGCGGACCCGATTCGGGCTTGGCCGATCTGTTT 60
DB 2755 ATGTTTGTAGAACATTACCGCCGCTCTGCGGACCCGATTCGGGCTTGGCCGATCTGTTT 2696
QY 61 CGTGCGGATGAACGTCCTCCGCGAATAATTAACTCCGGATTGGTGCTCTATAAAGATGAGACG 120
DB 2695 CGTGCGGATGAACGTCCTCCGCGAATAATTAACTCCGGATTGGTGCTCTATAAAGATGAGACG 2636
QY 121 GGCAAAACCCCGGTACTACACGCGTGAAGAGGCTGAACAGTATCTGCTCGAAAAATGAA 180
DB 2635 GGCAAAACCCCGGTACTACACGCGTGAAGAGGCTGAACAGTATCTGCTCGAAAAATGAA 2576
QY 181 ACCACAAAATAATTACTCTGGCATTTGACGGCATCCCTGAAATTTGGTCTGCTGCACTCAGGAA 240
DB 2575 ACCACAAAATAATTACTCTGGCATTTGACGGCATCCCTGAAATTTGGTCTGCTGCACTCAGGAA 2516

```
Qy 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAAAGTGTGCGACGGCAGACT 300
Db 2515 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAAAGTGTGCGACGGCAGACT 2456
Qy 301 CCGGGGGGCACTGGGGCACTACGGTGGCTGGCTCCGATTTCCCTGGCAAAAATACGAGGTT 360
Db 2455 CCGGGGGGCACTGGGGCACTACGGTGGCTGGCTCCGATTTCCCTGGCAAAAATACGAGGTT 2396
Qy 361 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGGCAACCAATAGAGCGCTCTTTAACTCTGCA 420
Db 2395 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGGCAACCAATAGAGCGCTCTTTAACTCTGCA 2336
Qy 421 GGTCTGGAAGTTCGTGAATACGCTTATATGATCGGGAATCACTCTTTGACTTCGAT 480
Db 2335 GGTCTGGAAGTTCGTGAATACGCTTATATGATCGGGAATCACTCTTTGACTTCGAT 2276
Qy 481 GCACGTATTACAGCCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTCCATGGCTGC 540
Db 2275 GCACGTATTACAGCCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTCCATGGCTGC 2216
Qy 541 TGCATTAACCCAAACCGGTATCGACCTACGCTGGAACAATGGCAAAACACTGGCAAACTC 600
Db 2215 TGCATTAACCCAAACCGGTATCGACCTACGCTGGAACAATGGCAAAACACTGGCAAACTC 2156
Qy 601 TCCGTTGAGAAAGGCTGGTTACCGCTGTTGACTTTCGCTTACGAGGGTTTTGCCCGTGGT 660
Db 2155 TCCGTTGAGAAAGGCTGGTTACCGCTGTTGACTTTCGCTTACGAGGGTTTTGCCCGTGGT 2096
Qy 661 CTGGAAGAAGTGTCTGAAGGACTGGCGGCTTTTCGGGCTATGCAATAAGAGCTGATGTT 720
Db 2095 CTGGAAGAAGTGTCTGAAGGACTGGCGGCTTTTCGGGCTATGCAATAAGAGCTGATGTT 2036
Qy 721 GCCAGTTCCTACTCTTAAACCTTTGGCTGTGTAACGAGCGTGTGGCGTTGTAACCTG 780
Db 2035 GCCAGTTCCTACTCTTAAACCTTTGGCTGTGTAACGAGCGTGTGGCGTTGTAACCTG 1976
Qy 781 GTTGCTGCCGACAGTGAACCGTTGATCGCGCAATTACGCCAAATGAAAGCGCGCATTCGC 840
Db 1975 GTTGCTGCCGACAGTGAACCGTTGATCGCGCAATTACGCCAAATGAAAGCGCGCATTCGC 1916
Qy 841 GCTAACTACTCTAAACCCACGAGCACAAGCGGCTTCTGTTGTTGCCACCATCCTGAGCAAC 900
Db 1915 GCTAACTACTCTAAACCCACGAGCACAAGCGGCTTCTGTTGTTGCCACCATCCTGAGCAAC 1856
Qy 901 GATCGGTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT 960
Db 1855 GATCGGTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT 1796
Qy 961 ATGCGTCAGTTGTTTCGTCAATACGCTGCAGGAAAAGCGCAACCCGCGACTTCAGCTTT 1020
Db 1795 ATGCGTCAGTTGTTTCGTCAATACGCTGCAGGAAAAGCGCAACCCGCGACTTCAGCTTT 1736
Qy 1021 ATCATCAACAGAAAGCGCATGTTCTCCTTCAGTGGCCCTGACAAAAGAACAAAGTCTCGT 1080
Db 1735 ATCATCAACAGAAAGCGCATGTTCTCCTTCAGTGGCCCTGACAAAAGAACAAAGTCTCGT 1676
Qy 1081 CTGCGCAAGAGTTTGGGATATAGCGGTTCGCTTCGCTTCGCTGCGTAAATGTGGCCGGATG 1140
Db 1675 CTGCGCAAGAGTTTGGGATATAGCGGTTCGCTTCGCTTCGCTGCGTAAATGTGGCCGGATG 1616
Qy 1141 ACACCAAGATACATGGCTCCGCTGCGAAGCGATTGGCGAGTGTGTA 1191
Db 1615 ACACCAAGATACATGGCTCCGCTGCGAAGCGATTGGCGAGTGTGTA 1565
```

RESULT 7

AAN71108

ID AAN71108 standard; DNA; 1293 BP.

XX AAN71108;

XX AC

XX XX

DT 01-JAN-1980 (first entry)

XX

DE Optimized Escherichia coli aspC gene.
XX aspC gene; feedback inhibition; amino acid synthesis; composite plasmid;
KW 86.
XX Escherichia coli.
OS
XX W08700202-A.
FN
XX 15-JAN-1987.
PD
XX 24-JUN-1986; 86WO-US001353.
PF
XX 24-JUN-1985; 85US-00747732.
PR
XX (NUTR-) NUTRASWEET CO.
PA
XX Edwards MR, Taylor PP, Hunter MG, Fotheringh IG;
PI WPI; 1987-021998/03.
XX P-PSDB; AAP70751.
DR
XX Composite plasmids contg. multiple genes in transcriptional units -
FT useful for prodn. of aminoacid(s), esp. L-phenylalanine and L-tyrosine.
FT
XX Disclosure; Page 25; 57pp; English.
XX
XX This sequence may be inserted into a composite plasmid and used for the
CC production of amino acids. See also AAN71053-55, AAN71107, AAN71109-11
CC and AAP70696-97 and AAP70750-54
XX
SQ Sequence 1293 BP; 322 A; 331 C; 343 G; 297 T; 0 U; 0 Other;

Query Match 99.7%; Score 1187.8; DB 1; Length 1293;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1189; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy 1 ATGTTTTCAGAACATTACCGCGCTCCTGCGACCCGATTTCTGGCCCTGGCCGATCTGTTT 60
Db 12 ATGTTTTCAGAACATTACCGCGCTCCTGCGACCCGATTTCTGGCCCTGGCCGATCTGTTT 71
Qy 61 CGTGCCGATGAAAGTCTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAGATGAGACG 120
Db 72 CGTGCCGATGAAAGTCTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAGATGAGACG 131
Qy 121 GGCACAAACCCCGGTACTGACCCAGCGTGAAAAGCTGAAACAGTATCTGCTCGAAAAATGAA 180
Db 132 GGCACAAACCCCGGTACTGACCCAGCGTGAAAAGCTGAAACAGTATCTGCTCGAAAAATGAA 191
Qy 181 ACCACCAAAATTAACCTCGGCAATTGACGGCATCCCTGAATTTGGTCCGTCACCTCAGGAA 240
Db 192 ACCACCAAAATTAACCTCGGCAATTGACGGCATCCCTGAATTTGGTCCGTCACCTCAGGAA 251
Qy 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAAAGTGTCTCGACGGCAGACT 300
Db 252 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAAAGTGTCTCGACGGCAGACT 311
Qy 301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTGCAAAAAATATACCAGCGTT 360
Db 312 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTGCAAAAAATATACCAGCGTT 371
Qy 361 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGGCAACCAATAGAGCGTCTTTAACTCTGCA 420
Db 372 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGGCAACCAATAGAGCGTCTTTAACTCTGCA 431
Qy 421 GGTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAAATCACACTCTTGACTTCGAT 480
Db 432 GGTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAAATCACACTCTTGACTTCGAT 491
Qy 481 GCACTGATTAAACAGCCTGAATGAAAGCTCAGGCTGGCGACGCTAGTGTGTTCCATGGCTGC 540
Db 492 GCACTGATTAAACAGCCTCAATGAAGCTCAGGCTGGCGACGCTAGTGTGTTCCATGGCTGC 551
```

QY 541 TGCATAACCCAAACGGTATCGACCTTACGCTGGAACAATGGCAACATCGGCACAACTC 600
DB 552 TGCATAAACCCAAACGGTATCGACCTTACGCTGGAACAATGGCAACATCGGCACAACTC 611
QY 601 TCCGTTGAGAAAGGCTGGTTACCGCTGTTTGAAGCTTACCGCTTACCGAGGTTTGGCCGTTGGT 660
DB 612 TCCGTTGAGAAAGGCTGGTTACCGCTGTTTGAAGCTTACCGCTTACCGAGGTTTGGCCGTTGGT 671
QY 661 CTGGAAGAAGATGCTGAAGACTGCGCGCTTTCGGGCTATGATGAAGAAGCTGATGTT 720
DB 672 CTGGAAGAAGATGCTGAAGACTGCGCGCTTTCGGGCTATGATGAAGAAGCTGATGTT 731
QY 721 GCCAGTTCCTACTCTAAACCTTTGGCTCTGACAGAGCTGTGCGCTTCTGACTCTG 780
DB 732 GCCAGTTCCTACTCTAAACCTTTGGCTCTGACAGAGCTGTGCGCTTCTGACTCTG 791
QY 781 GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTTCG 840
DB 792 GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTTCG 851
QY 841 GCTAACTACTCTAACCCACAGCAGCAGCGGCTTCTGTTGTCACCATCTCTGAGCAAC 900
DB 852 GCTAACTACTCTAACCCACAGCAGCAGCGGCTTCTGTTGTCACCATCTCTGAGCAAC 911
QY 901 GATGCGTTACGTGGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT 960
DB 912 GATGCGTTACGTGGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT 971
QY 961 ATGCGTCAGTTGTCGTCATAGCTGCAGGAAAAAGCGCAACACCGCACTTTCAGCTTT 1020
DB 972 AGCGTCAGTTGTCGTCATAGCTGCAGGAAAAAGCGCAACACCGCACTTTCAGCTTT 1031
QY 1021 ATCATCAACAGAACCGCATGTTCTCCTCAGTGGCTGACAAAGAACAAAGTGTGCGT 1080
DB 1032 ATCATCAACAGAACCGCATGTTCTCCTCAGTGGCTGACAAAGAACAAAGTGTGCGT 1091
QY 1081 CTGCGCAAGAGTTGCGGTATATGCGGTTGCTTCTGCTCGCGTAAATGTTGGCCGGATG 1140
DB 1092 CTGCGCAAGAGTTGCGGTATATGCGGTTGCTTCTGCTCGCGTAAATGTTGGCCGGATG 1151
QY 1141 ACACGATTAACATGCTCGCTGTGGAAGCGATTTGGCAGTGTCTGTAA 1191
DB 1152 ACACGATTAACATGCTCGCTGTGGAAGCGATTTGGCAGTGTCTGTAA 1202

RESULT 8
ACH99107
ID ACH99107 standard; DNA; 1209 BP.
AC ACH99107;
XX
DT 29-JUL-2004 (first entry)
XX
DE Klebsiella pneumoniae polynucleotide seqid 4902.
XX
KW Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.
XX
OS Klebsiella pneumoniae.
XX
FN US6610836-B1.
XX
PD 26-AUG-2003.
XX
PF 27-JAN-2000; 2000US-00489039.
XX
PR 29-JAN-1999; 99US-0117747P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL, Osborne M;
XX
XX WPI; 2003-895346/82.
DB

DR P-SDB; ABO65556.
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
XX
PS Disclosure; SEQ ID NO 4902; 932pp; English.
XX
CC The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
CC pneumoniae polypeptide of the invention
XX
SQ Sequence 1209 BP; 271 A; 353 C; 343 G; 242 T; 0 U; 0 Other;

Query Match 70.4%; Score 839; DB 11; Length 1209;
Best Local Similarity 81.5%; Pred. No. 1.8e-256;
Matches 971; Conservative 0; Mismatches 220; Indels 0; Gaps 0;

QY 1 ATGTTTGAGAACATTACCGCCCTCTGCGGACCCGATTCGGCCCTGGCGATCTGTTT 60
DB 19 ATGTTTGAGAACATTACCGCCCTCTGCGGACCCGATTCGGCCCTGGCGATCTGTTT 78
QY 61 CGTGGCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTCTCTATAAAGATGAGACG 120
DB 79 CGTGGCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTCTCTATAAAGATGAGACG 138
QY 121 GCGAAACCCCGGATCTGACACGCGTGAAGAGGCTGAACAGTATCTGCTCGGAAATGAA 180
DB 139 GGTAAACCGCCCTTCTGACACGCGTCAAAAAAGCAGAGCAGTATCTGCTGGAATGAA 198
QY 181 ACCACCAAAATTAACCTCGGATTTGACGCGATCCCTGAATTTGGTCTGCTCAGGAA 240
DB 199 AGCACTAAAAACATATCTGGGCTATCGATGATTTCTCTGAATTTGGTCTGCAACCCAGGAG 258
QY 241 CTGCTGTTTGTAAAGTAGCGCCCTGATCAATGACAAACGCTGCTCGCACGGCACAGACT 300
DB 259 CTGCTGTTTGTAAAGTAGCGCCCTGATCAATGATGATGCGGAAATCAACACTTTGACTTCGAT 318
QY 301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGCAAAAAATACCAGCTT 360
DB 319 CCGGGGGTACCGGTGCGCTGCTGCGCGAGACTTCTCTGCGCAAAAAACACCGAGCTG 378
QY 361 AAGCGTGTGGGTGAGCAACCCAGCTGCGGCAACATAGAGCGCTTTTAACCTTGCA 420
DB 379 AAAACGTGTGGGTAAAGTAAATCCGAGCTGGCGCAACCATANAAGCGTATTTACCTCTGCC 438
QY 421 GGTCTGGAAGTTGCTGAATACGCTTATATGATGCGGAAATCAACACTCTTGACTTCGAT 480
DB 439 GGGCTGGAAGTTGCGGAATACGCTTACTAGACCGCGCTTAACCAACGGCTTGGACTTTGAT 498
QY 481 GCACTGATTAAACAGCCCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTTCCATGGCTGC 540
DB 499 GGTCTGCTGCCAGCTTGAACGAGGCCAGCGGGGCGAGCTGTACTGTTTCCACGGCTGC 558
QY 541 TGCATPAACCCCAACCGGTATCGACCTTACGCTGGAAACAATGGCAAAACACTGGCACAACTC 600
DB 559 TGCACAACCCGACCGGTATCGATCCGAGCTCGATCAGTGGCGAGCAGCTGGCGCAGCTG 618
QY 601 TCCGTTGAGAAAGGCTGGTTTACCGCTTGTGACTTACCGGTTTACCAGGTTTGGCCGTTGGT 660
DB 619 TCCGTTGAGAAAGGCTGGCTTACCGCTGTTTCGATTTTCGCTTACCAGGCTTCCGCCCGGT 678
QY 661 CTGGAAGAAGATGCTGAAGACTTGGCGCTTTTCGGGCTATGATGAAGAAGCTGATGTTT 720
DB 679 CTGGAAGAAGATGCTGAAGGCTTACCGGCTTTTGTGCTCTCAATAAAGAGCTGCTGCTC 738
QY 721 GCCAGTTCCTACTCTAAAAAATTTTGGCCTGTGAACAGAGCGTGTGGCGCTTGTACTCTG 780
DB 739 GCCAGTTCCTACTCTAAAAAATTTTGGCCTGTGAACAGAGCGCTGCGCGCTCTGCACTCTG 798

Qy 781 GTTCTGCGCAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGGATTCGC 840
 Db 799 GTCCCGCGGATCAGGAGACTGTAGACCGCCCTTCAGTCAGATGAAGTGGTGCATCGC 858
 Qy 841 GCTAACTACTTAACCCACAGACACAGCGCGCTTCTGTTGTTGCCACCATCTGAGCAAC 900
 Db 859 GCCAACTACTCGAACCAGCGCTGCGCATGGCGCCTCCGTGGTTCGCCACCATTCAGCAAC 918
 Qy 901 GATCGGTTACGTGCGATTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960
 Db 919 GATCGGTCAGCGCAATCTGGAGCAGGAAGTACCGATATGCGCCAGCGCATCCAGGT 978
 Qy 961 ATGCGTCAGTTGTTGCTCAATACCTGCGAGAAAAGCGCAACCCGACCTTCAGCTTT 1020
 Db 979 ATGCGTCTGCTGTTGCTCAATACCTGCGAGGAGAAAGCGCGAGCTTCAGCTTT 1038
 Qy 1021 ATCATCAACAGACGCGATGTTCTCTTCTAGTGGCTGACAAAGAACAGTCTCGGT 1080
 Db 1039 ATCAGCCAGCAGAAACGCGATGTTCTCATTTACGCGCCCTGACTAAAGAGCAGGTGCTCGC 1098
 Qy 1081 CTGCGCGAAGAGTTGGCGTATATGCGTTGCTTCTGTCGCGTAAATGTGGCCGGGATG 1140
 Db 1099 CTGCGTGAAGATTCGCGCATCTATGCGGTAGCTTCCGAGCATATCAACGTGGCGGGATG 1158
 Qy 1141 ACACGAGTAACATGGCTCCGCTGTGCGAAGCGATTGTGGCAGTGTCTGTAA 1191
 Db 1159 ACGCTGACATATGGCGCGCTGTGCGAGCCATCTGCGCGTACTGTAA 1209

RESULT 9

ACF68649
 ID ACF68649 standard; DNA; 1191 BP.
 XX ACF68649;
 AC ACF68649;
 DT 20-NOV-2003 (first entry)
 XX Photorhabdus luminescens nucleotide sequence #7116.
 DE Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough; gene; ds.
 XX Photorhabdus luminescens.
 OS Photorhabdus luminescens.
 XX W0200294867-A2.
 PN 28-NOV-2002.
 XX 07-FEB-2002; 2002WO-IB003040.
 XX 07-FEB-2001; 2001FR-00001659.
 XX (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;
 XX WPI; 2003-148459/14.
 XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX Claim 2; SEQ ID NO 7116; 1205pp; French.
 XX The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification

CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens genes
 XX

SQ Sequence 1191 BP; 346 A; 279 C; 267 G; 299 T; 0 U; 0 Other;
 Query Match 58.1%; Score 691.8; DB 10; Length 1191;
 Best Local Similarity 73.8%; Pred. No. 1.7e-209;
 Matches 879; Conservative 0; Mismatches 312; Indels 0; Gaps 0;

Qy 1 ATGTTTGAGAACATTACCGCGCTCTCGCGACCCGATTTGCGCCCTGGCCGATCTGTTT 60
 Db 1 ATGTTTGAGAAAATCACCGCAGCGCTCGCGACCCCTATTCTTGGCTTAGCCGATAGTTTC 60
 Qy 61 CGTGGCGATGAACGTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAGATGAGACG 120
 Db 61 CGTTCGTGATCCTGTACAAATTAATCACTTTGGTATCGGTGCTATTAAGAGCAAAACA 120
 Qy 121 GGCAAAACCCCGGTACTGACGCGTGAAAAGGCTGAACAGTATCTGCTCGAAAATGAA 180
 Db 121 GGAAAACCCCGGTCTTGACAGTGTTTAAAAAGCTGAACAATATTACTTGAAGAAACGA 180
 Qy 181 ACCACAAAATTTACTCTGGCATGACGGCATCCCTGAATTTGTCGCTGCACACTCAGGAA 240
 Db 181 ACAACAAAGAAATTTACTGCGGCTTAGCGGCTTAGCGGAATTTGCGCCGCTAACTCAAGAA 240
 Qy 241 CTGCTGTTTGTAAAGTAGCGCCCTGATCAATGACAAACGTCTCCACGCGCACAGACT 300
 Db 241 TTACTGTTTGGCAAGATACCCAGTTGTACAGATTAACGCGCCCGACAGCAAAAGC 300
 Qy 301 CCGGGGGCGACTGCGCGCATACGCGTGGCTGCGGATTTCTCTGGCAAAAATACCAGCGTT 360
 Db 301 CCAGCGGTACTCGGTGCTTTACGTATTGCTGCGGATTTTATTCACCAACAGACTAATGCT 360
 Qy 361 AAGCGTGTGGGTGAGCAACCCAGGCTGGCGAACCATAAGACGCTCTTTAACTCTGCA 420
 Db 361 AAACGAGTTTGGATCAGCAACCCCAACCTGGCCAAACCATAAAAACGTTTTTCCGCGCT 420
 Qy 421 GGTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAATCACACTCTTGACTTCGAT 480
 Db 421 GGTCTGGAAGTTCGTGAATATAAATCTATGACGCTGAAAACACGCGCTGAATTTCCGA 480
 Qy 481 GCACTGATTAAACAGCCCTGAATGAAGCTCAGGCTGGCGACGCTAGTGTGTTCCATGGCTGC 540
 Db 481 GACATGCTGCAAGCGCTGTCGAAAGCTCAGCGTGGTGTGTTGTTCTGTTCCACGGCTGC 540
 Qy 541 TGCATTAACCCAAACCGGTATCGACCCCTACGCTGGAAACAATGGGCAACACCTGGCAACATC 600
 Db 541 TGCACAAATCCGACAGGCATCGATCCAAACCCCGCACAAATGGGCTAAATGGCAGAAATG 600
 Qy 601 TCCGTTTGAGAAAGGCTGTTACCGCTGTTTGAATCTTCCCTTACCAGGCTTTTCCCGGTGT 660
 Db 601 TCTCGGAGAAAGGCTGTTGCTTATTTTGAATTTTCCCTTACCAGGATTTTCCCAAGGCG 660
 Qy 661 CTGCAAGAGATGCTGAAGGACTGCGCGCTTTTCGCGCTATGCAATAAAGAGCTGATGTT 720
 Db 661 CTAAACGAAGATGAGAGGCGCTTACGTATTTTTCGGAATAATCATATGAATGATGTT 720
 Qy 721 GCCAGTTCTTACTCTAAAAAATTGGCGCTGTACAAACGAGCGTGTGGCGCTGTGACTGT 780
 Db 721 GCCAGCTCTTACTCCAAAACCTTTGGCGCTGTACAAATGAACGCTGTGCGTGCCTGTACTATT 780

Qy	781	GTTCCTGCCGACAGTGAACCGTTGATCTCGGCATTCAGCCAAATGAAGCGGCATTCGC	840
Db	781	GTTCCTAGTGACAGTGATACAGCAGAAAAGCGTTGAGCCACAGTAAGCGATATCCGT	840
Qy	841	GCTAACTACTCTAAACCCACCAAGCAGCAGCGCGCTTCTGTTGTGCCACCATCTTGAGCAAC	900
Db	841	GCTAACTATTCCAAACCCACCGCTCATGGTGCATCTATTGTCACCTACCAATTTGTCAAT	900
Qy	901	GATCGGTTACGTGCGATTTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT	960
Db	901	GAAGACCTGAAAGCAGCTTGGGAACAGGAACTGACCAACCATGCGCGAACTATCCAGCGT	960
Qy	961	ATGCGTCAGTTGTTGCTCAATACGCTGCAGGAAAAAGCGGCAACCCGACATTCAGCTTT	1020
Db	961	ATGCGTCAACTGTTTGTCAATACCTTGCAGGAAAAGGGGCNAAACAGGATTTTCAGCTTT	1020
Qy	1021	ATCATCAACAGACGCGCATGTTCTCTTCAGTGGCCCTGCACAAAGAACAAAGTCTCSCGT	1080
Db	1021	ATTATTAGCCAAAATGGTATGTTCTCATTCAGTGGCCCTGCACAAAGAACAAAGTAGAACGT	1080
Qy	1081	CTGCGCGAAGAGTTTGGCGGTATATGCGGTGCTTTCTGGTCGCGTAAATGTGGCGCGGATG	1140
Db	1081	CTGGGTGATGAGTTTGGTATATACGCTGTGAGTTCGCGTCGTATTACGTCGCTGGCTTG	1140
Qy	1141	ACACGAGATAACATGGCTCCGCTGTGCGAAGCGAATTGTGGCAGTGTGTAA	1191
Db	1141	ACGTTAGAAAACATGGCTCCACTATGTGAAGCCATTGTTGCAGTACTCTAA	1191

WP	Fragment Name	Begin	End
WP	ACF67367_00	1	110000
WP	ACF67367_01	100001	210000
WP	ACF67367_02	200001	310000
WP	ACF67367_03	300001	410000
WP	ACF67367_04	400001	510000
WP	ACF67367_05	500001	610000
WP	ACF67367_06	600001	710000
WP	ACF67367_07	700001	810000
WP	ACF67367_08	800001	910000
WP	ACF67367_09	900001	1010000
WP	ACF67367_10	1000001	1110000
WP	ACF67367_11	1100001	1210000
WP	ACF67367_12	1200001	1310000
WP	ACF67367_13	1300001	1410000
WP	ACF67367_14	1400001	1510000
WP	ACF67367_15	1500001	1610000
WP	ACF67367_16	1600001	1710000
WP	ACF67367_17	1700001	1810000
WP	ACF67367_18	1800001	1910000
WP	ACF67367_19	1900001	2010000
WP	ACF67367_20	2000001	2110000
WP	ACF67367_21	2100001	2210000
WP	ACF67367_22	2200001	2310000
WP	ACF67367_23	2300001	2410000
WP	ACF67367_24	2400001	2510000
WP	ACF67367_25	2500001	2610000
WP	ACF67367_26	2600001	2710000
WP	ACF67367_27	2700001	2810000
WP	ACF67367_28	2800001	2910000
WP	ACF67367_29	2900001	3010000
WP	ACF67367_30	3000001	3110000
WP	ACF67367_31	3100001	3210000
WP	ACF67367_32	3200001	3310000
WP	ACF67367_33	3300001	3410000
WP	ACF67367_34	3400001	3510000
WP	ACF67367_35	3500001	3610000
WP	ACF67367_36	3600001	3710000
WP	ACF67367_37	3700001	3810000

WP	ACF67367_38	3800001	3910000	
WP	ACF67367_39	4000001	4010000	
WP	ACF67367_40	4000001	4110000	
WP	ACF67367_41	4100001	4210000	
WP	ACF67367_42	4200001	4310000	
WP	ACF67367_43	4300001	4410000	
WP	ACF67367_44	4400001	4510000	
WP	ACF67367_45	4500001	4610000	
WP	ACF67367_46	4600001	4710000	
WP	ACF67367_47	4700001	4810000	
WP	ACF67367_48	4800001	4910000	
WP	ACF67367_49	4900001	5010000	
WP	ACF67367_50	5000001	5110000	
WP	ACF67367_51	5100001	5210000	
WP	ACF67367_52	5200001	5310000	
WP	ACF67367_53	5300001	5410000	
WP	ACF67367_54	5400001	5510000	
WP	ACF67367_55	5500001	5610000	
WP	ACF67367_56	5600001	5648894	

Query Match

Best Local Similarity

Matches 879; Conservative

58.1%;

73.8%;

0; Mismatches 312; Indels 0; Gaps 0;

Score 691.8; DB 10; Length 110000;

Pred. No. 2.1e-208;

Qy	1	ATGTTTGAGAACTATACCGCGCTCTCCGCGACCGGATCTTGGCGCTGGCCGATCTGTTT	60
Db	23013	ATGTTTGAGAAAATCACCGACGCGCTCCGCAACCCTATTCTTGGCTTAGCCGATGTTTC	23072
Qy	61	CGTGGCGATGAACGTCCTGGGCAAAATTAACCTCGGATTTGGTGCTATAAAGATGAGACG	120
Db	23073	CGTTCTGATCTCGTACAAAATAAAATCAACCTTGGTATCGGTGTCATAAAGACGAACA	23132
Qy	121	GGCAAAACCCCGGTACTGACCAGCGTGAAGAAGGCTGAACAGTATCTGCTCGAAAAATGAA	180
Db	23133	GGAAAAACCCCGGTTCTGACCAAGTGTAAANAAGCTGAACAATATTACTGGAACAAGAA	23192
Qy	181	ACCACCAAAATTTACCTCGGCAATGACGGCATTCCTCGAATTTGGTCGCTGCACTCAGGAA	240
Db	23193	ACAAACAAGAATTTATCTGCCGATTAGCGGCTTAGCCGAATTTGGCCGCTTAACCTCAAGAA	23252
Qy	241	CTGCTGTTTGGTAAAGGTAGCCCTGATCAATGACAAACGTGCTCGCACGGCACACAGCT	300
Db	23253	TTACTGTTTGGCAAGATCACCCAGTTGTACAGATAAACCGCCGCCACAGCAACAAGC	23312
Qy	301	CGGGGGGCACCTGGCGCACCTACGCGTGCGTCCGATTTCTTGGCAAAAAATACCAGCGTT	360
Db	23313	CCAGGCGGTACCGGTGCTTTAGTATTGCTCGGATTTCAATGGCAACAGCTAATGCT	23372
Qy	361	AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGGCAACCAATAGAGGCTTTTAACCTGCA	420
Db	23373	AAACGAGTTTGGATCAGCAACCCAAACCTGGCCAAACCAATAAAAAACGTTTTTCCGCGCT	23432
Qy	421	GGTCTGGAAGTTTCGTGMAATACGCTTATTATGATCGGAAATACACACTCTTGTACTCGAT	480
Db	23433	GGTCTGGAAGTCTGTGAATATAAATFATGACGCTGANAACACGGCTGAATTTGCA	23492
Qy	481	GCACCTGATTAAACAGCCTGAAATGAAGCTCAGGCTGGCGACGTAGTGTCTTCCATGGCTGC	540
Db	23493	GACATGCTGGCAAGCCTGTCGAAAGCTCAGGCTGGTGATGTTGTTCTGTTCCACGGCTGC	23552
Qy	541	TGCCATAACCCAAACCGGTATGCAACCTTACGCTGGAAACAATGGCAAAACACTGGCAACATC	600
Db	23553	TGCCAACAATCCGACAGGCATCGATCCAAACCCCGGCACAATGGGCTTAAACTGGCAGAAATG	23612
Qy	601	TCCGTTGCAAAAGGCTGGTTACCGCTGTTTGACTTTCGCTTACCAAGGTTTTTGGCCGTGGT	660
Db	23613	TCTGCGGAAGAAAGGCTGGTTGCTATTTTTGATTTTCGCTTACCAAGGATTTTGCAAGGGC	23672
Qy	661	CTGGAAGAAAGATGCTGAAGGACTGGCGGCTTTTCGCGGCTATGCATATAAAGAGCTGATGTT	720
Db	23673	CTAACAAGAGATGCAGAGGGCCCTACGTATTTTGGCAAAATCATATATGAATGATGTT	23732
Qy	721	GCAGTTCTCTACTCTATAAAACCTTTGGCGCTGTACAACAGAGCGTGTGGCGCTGTGACTCTG	780

Db 23733 GCCAGCTCTTACTCAAAATTTGGGCTGTACAAATGAAAGTGGCTGCTGTACTATT 23792
 Qy 781 GTTGCTGCCGACAGTGAACCGTTGATCGGCGCATTCAGCCAAATGAAAGCGGCGATTCGC 840
 Db 23793 GTTGCTAGTACAGTGATACAGCAGAAAAGCGTTCAGCCAAAGCTAAAGCGATTATCCGT 23852
 Qy 841 GCTAACTACTTAAACCCAGCAGACACGGCGCTTCTGTTGTTGCCACCTCCTGAGCAAC 900
 Db 23853 GCTAACTATTCCAAACCCAGCGCTCATGCTGCTATTTGCTACTACCAATTTGTCAAT 23912
 Qy 901 GATCGCTAGCTGCGATTGGGACACAGAGCTGACTGATATGCGCCAGCGTATTCACCGT 960
 Db 23913 GAAGACCTGAAGCAGCTTGGGAACAGGAACTGACCCACATGCGCGAACGTATCCAGCGT 23972
 Qy 961 ATGCGTCAGTTGTTGCTCAATACGCTCGAGAAAAGGCGCAACCGCGACTTTCAGCTTT 1020
 Db 23973 ATGCGTCAGTTGTTGCTCAATACCTGACGAAAAGGCGCAACAGGATTTTCAGCTTT 24032
 Qy 1021 ATCATCAACAGAAACCGCATGTTCTCTCAGTGGCTGACAAAGAACCAAGTCTGCGT 1080
 Db 24033 ATTATTAGCCAAATGATGTTCTCTATTGAGTGGCTGACAAAGAACCAAGTAGAACGT 24092
 Qy 1081 CTGCGGCAAGATTTGCGGTATATGCGGTGCTTCTGCTGCGTAAATGTGGCGGGATG 1140
 Db 24093 CTGCGTGATGAGTTGGTATATACGCTGTGAGTTCCGCTGCTATTAACGTCGCTGGCTG 24152
 Qy 1141 ACACCATCAACATGCTCGCTGTGCGAAGCGATTGTGCGAGTGTGTTAA 1191
 Db 24153 ACGTTAGAAACATGGCTCCACTATGTGAAGCCATTGTTGAGTACTCTAA 24203

RESULT 11

ACF65381
 ID ACF65381 standard; DNA; 249878 BP.
 XX
 AC ACF65381;
 DT 20-NOV-2003 (first entry)
 XX
 DE Photorhabdus luminescens nucleotide sequence #34.
 XX
 KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough; gene; ds.
 XX
 OS Photorhabdus luminescens.
 XX
 WO200294867-A2.
 XX
 PD 28-NOV-2002.
 XX
 PF 07-FEB-2002; 2002WO-IB003040.
 XX
 PR 07-FEB-2001; 2001PR-00001659.
 XX
 PA (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;
 XX
 DR WPI; 2003-148459/14.
 XX
 PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX
 PS Claim 1; SEQ ID NO 34; 1205pp; French.
 XX
 CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens

CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens genes
 XX
 SQ Sequence 249878 BP; 75946 A; 56552 C; 49249 G; 68029 T; 0 U; 2 Other;
 Query Match 58.1%; Score 691.8; DB 10; Length 249878;
 Best Local Similarity 73.8%; Pred. No. 3.3e-208;
 Matches 879; Conservative 0; Mismatches 312; Indels 0; Gaps 0;
 Qy 1 ATGTTTGGAGAACATTACCGCGCTCTCTGCGACCCGATTTCTGGGCGCTGGCGGATCTCTTT 60
 Db 1334 ATGTTTGGAGAAATACCGCAGCGCTGCGACCCCTATTCTTGGCTTAGCGGATGTTTC 1393
 Qy 61 CGTGCCGATGAACGTCCTCGGCAAAATTAACCTCGGATGCTGCTATAAGATGAGACG 120
 Db 1394 CGTTCGTATCTCGTACAAATAAATCAACCTTGGTATCGGTCTCTATAAGACGAAACA 1453
 Qy 121 GGCAAAACCCGCTACTGACCGCTGAAAAGGCTGAACAGTATCTGTCGGAATGAA 180
 Db 1454 GGAAAAACCCCGGTTCTGACCAAGTGTAAAAAGCTGAACAATATTACTTGGAAAAACGAA 1513
 Qy 181 ACCACCAAAAAATTACCTCGGCATTGACGGCATCCCTGAAATTTGGTCTGCTGCACCTCAGGAA 240
 Db 1514 ACAACAAGAAATATTCTGCCGATTAGCGGCTTAGCCGAATTTGCGCGGTAACCTCAAGAA 1573
 Qy 241 CTGCTCTTTGTTAAAGTAGCGCCCTGATCAATGACAAAGCTGCTCGCAGGACAGACT 300
 Db 1574 TTACTGTTTGGCAAGATCACCCAGTTGTACAGATAAACCGCCCGCAGCAGACAAAGC 1633
 Qy 301 CCGGGGGCACTGGCGCACTAGCGTGGCTGCGGATTTCTTGGCAAAAAATACCGAGTT 360
 Db 1634 CCAGGCGGTACCGGTGCTTTAGCTATTGCTGCGGATTTTCATTCGCAACACAGACTAATGCT 1693
 Qy 361 AAGCGTGTGTGGGTGAGCAACCCCAAGCTGCGCCGAACCAATGAAGAGCGTCTTTAACTCTGCA 420
 Db 1694 AACAGATTTGGATCAGCAACCCNACCTGGCCAAACATATAAAACGTTTTTTCGCGCGCT 1753
 Qy 421 GGTCTGGAAGTTGTTGAATACGCTTATATGATGCGGAAATACACTCTTTGACTTCGAT 480
 Db 1754 GGTCTGGAAGTCTGTAATATAAATACTATGACGCTGAAAAACACGCGCTGAATTTTCGAA 1813
 Qy 481 GCACTGATTAAACGCTGTAATGAGCTCAGGCTGGCGACGCTAGTCTGTTTCCATGGCTGC 540
 Db 1814 GACATGCTGGCAAGCGCTTCCGAAAGCTCAGGCTGGGTGATGTTGTTCTGTTCCACGCGTGC 1873
 Qy 541 TGCATATAACCAACCGGTATCGACCTAGCTGGAACCAATGGAACACACCTGGGACCAACTC 600
 Db 1874 TGCCACATCCGACAGGCAATGATCCAAACCCGCGCAATGGGCTAACTGGCAGAAATG 1933
 Qy 601 TCGTTTGGAGAAAGGCTGGTTACCGCTGTTGACTTTCGCTTACAGGCTTTTCCCGGTGGT 660
 Db 1934 TCTCGGAGAAAGGCTGGTTGCTCTATTTTGTATTACCAAGGATTTGCAAGGCGC 1993
 Qy 661 CTGGAAGAAGATGCTGAAGACTGCGCGCTTTTCGCGGCTATGCTATAAGAGCTGATTGTT 720
 Db 1994 CTAACGAAGATGACAGAGGCGCTACGTTATTTTTCGAAAAAATCATTAATGAACCTGATTGT 2053

ID AAS90082 standard; cDNA; 3222 BP.
XX AC AAS90082;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #25886.
XX DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG25895.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
XX biodiversity.
PS Claim 1; SEQ ID NO 25886; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 3222 BP; 727 A; 880 C; 863 G; 752 T; 0 U; 0 Other;

Query Match 41.7%; Score 496.2; DB 5; Length 3222;
Best Local Similarity 99.4%; Pred. No. 8.2e-147;
Matches 498; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 687 CGCTTTCCGGCTATGCATAAAGAGCTGATGTTGCCAGTTCCTACTCTAAAACTTTGG 746
DB 1425 CCCTTTCCGGCTATGCATAAAGAGCTGATGTTGCCAGTTCCTACTCTAAAACTTTGG 1484

QY 747 CCTGTACAACGACGCTTTGGCGCTTGTACTCTGTTGCTGCCAGAGTGAACCGTTGA 806
DB 1485 CCTGTACAACGACGCTTTGGCGCTTGTACTCTGTTGCTGCCAGAGTGAACCGTTGA 1544

QY 807 TCGCGCATTCAGCAAAATGAAAGCGGCGATTTCGCGCTAACTACTCTAAACCCAGCACA 866

Db 1545 TCGCGCATTCAGCAAAATGAAAGCGGCGATTTCGCGCTAACTACTCTAAACCCAGCACA 1604
QY 867 CGGCGCTTCTGTTGTTGGCCACCATCTGTAGCAACGATCGGTTACGTGCGATTTGGGAACA 926
Db 1605 CGGCGCTTCTGTTGTTGGCCACCATCTGTAGCAACGATCGGTTACGTGCGATTTGGGAACA 1664
QY 927 AGAGCTGACTGATATCGCCAGCGTATTTCAGCGTATGCGTCACTTGTTCGTCAATACGCT 986
Db 1665 AGAGCTGACTGATATCGCCAGCGTATTTCAGCGTATGCGTCACTTGTTCGTCAATACGCT 1724
QY 987 GCAGGAAAAAGCGGCAAAACCGGACCTTCAGCTTTATCATCAACAGAACGGCATGTTCTC 1046
Db 1725 GCATGAAAAAGCGGCAAAACCGGACCTTCAGCTTTATCATCAACAGAACGGCATGTTCTC 1784
QY 1047 CTTTCAGTGGCTTGACAAAAAGAACCAAGTGTCTGCGGAAAGAGTTTGGCGGTATATGC 1106
Db 1785 CTTTCAGTGGCTTGACAAAAAGAACCAAGTGTCTGCGGAAAGAGTTTGGCGGTATATGC 1844
QY 1107 GGTTCCTTCTGTCGCGTAAATGTGCGCGGATGACACAGATAAATGCTCCGCTGTG 1166
Db 1845 GGTTCCTTCTGTCGCGTAAATGTGCGCGGATGACACAGATAAATGCTCCGCTGTG 1904
QY 1167 CGAAGCGATTGCGCAGTGCT 1187
Db 1905 CGAAGCGATTGCGCAGTGCT 1925

Search completed: March 15, 2006, 03:02:48
Job time : 1397 secs

This Page Blank (uspto)

2344 AIGITTAGAACAATTACCGCGCTCTGTGCCACCGGATCTGGGCGCTGGCCGATCTGTTT(240)

43

Qy	61	CGTGCCGATGAAACGCTCCCGGCAAAATTAACCTCGGGATTTGGTGCTATATAAGATGAGACG	120
Db	2404	CGTGCCGATGAAACGCTCCCGGCAAAATTAACCTCGGGATTTGGTGCTATATAAGATGAGACG	2463
Qy	121	GGCAAAACCCGGTACTGACGACGGTGAAAAGGCTGAAACAGTATCTGCTCGAATAAGAA	180
Db	2464	GGCAAAACCCGGTACTGACGACGGTGAAAAGGCTGAAACAGTATCTGCTCGAATAAGAA	2523
Qy	181	ACCACCAAAATTAACCTCGGCATTTGACGGGATTCCTGAAATTTGGTGCCTGCACCTCAGGAA	240
Db	2524	ACCACCAAAATTAACCTCGGCATTTGACGGGATTCCTGAAATTTGGTGCCTGCACCTCAGGAA	2583
Qy	241	CTGCTGTTTGGTAAAGTAGCGCCTCATCAATGACAAACGTCGTCTGCGACGGACAGACT	300
Db	2584	CTGCTGTTTGGTAAAGTAGCGCCTCATCAATGACAAACGTCGTCTGCGACGGACAGACT	2643
Qy	301	CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGATTTCTTGGCAGAAAATATACACGGTT	360
Db	2644	CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGATTTCTTGGCAGAAAATATACACGGTT	2703
Qy	361	AAGCGTGTGTGGGTGAGCAACCCAAAGCTGGCGGCAACATAAGACGGTCTTTAACTCTGCA	420
Db	2704	AAGCGTGTGTGGGTGAGCAACCCAAAGCTGGCGGCAACATAAGACGGTCTTTAACTCTGCA	2763
Qy	421	GGTCTGGAAGTTTCGTGAATACGCTTATTATGATCGGAAAATCACTCTTGACTTCGAT	480
Db	2764	GGTCTGGAAGTTTCGTGAATACGCTTATTATGATCGGAAAATCACTCTTGACTTCGAT	2823
Qy	481	GCACGTGTTAACAGCCTGAATGAAAGCTCAGGCTGGGACGTAGTGCTGTTCATGGCTGC	540
Db	2824	GCACGTGTTAACAGCCTGAATGAAAGCTCAGGCTGGGACGTAGTGCTGTTCATGGCTGC	2883
Qy	541	TGCCATAACCCAAACCGGTATCGACCCTACGCTGGAAACAAATGGCAAAACACTGGGCACAACTC	600
Db	2884	TGCCATAACCCAAACCGGTATCGACCCTACGCTGGAAACAAATGGGCACAACTC	2943
Qy	601	TCCGTTGAGAAAGGCTGTTTACCCTGTTTCACTTCGCTTACCAGGGTTTTGCCCGTGGT	660
Db	2944	TCCGTTGAGAAAGGCTGTTTACCCTGTTTCACTTCGCTTACCAGGGTTTTGCCCGTGGT	3003
Qy	661	CTGGAAGAGATGCTGAAGGACTGGCGCTTTTCGGGCTATGCAATAAGAGCTGATTTGTT	720
Db	3004	CTGGAAGAGATGCTGAAGGACTGGCGCTTTTCGGGCTATGCAATAAGAGCTGATTTGTT	3063
Qy	721	GCCAGTTTCTTACTCTTAAAAACTTTTGGCTGTACAAACGAGCTGTTGGCGTTGTACTCTG	780
Db	3064	GCCAGTTTCTTACTCTTAAAAACTTTTGGCTGTACAAACGAGCTGTTGGCGTTGTACTCTG	3123
Qy	781	GTTCCTGCCGACAGTGAACCGTTGATGCGGCATTTACGCCAAATGAAGGGCGGATTCGC	840
Db	3124	GTTCCTGCCGACAGTGAACCGTTGATGCGGCATTTACGCCAAATGAAGGGCGGATTCGC	3183
Qy	841	GCTAACTACTTAACCCACGAGCAACGGCGCTTCTGTTTGGCCACCATCTCGAGCAAC	900
Db	3184	GCTAACTACTTAACCCACGAGCAACGGCGCTTCTGTTTGGCCACCATCTCGAGCAAC	3243
Qy	901	GATGCGTTTACGTGCGATTTTGGGAAACAGAGCTGACTGATATGCGCCAGCGTATTACGCGT	960
Db	3244	GATGCGTTTACGTGCGATTTTGGGAAACAGAGCTGACTGATATGCGCCAGCGTATTACGCGT	3303
Qy	961	ATGCGTCAGTTGTTTCGTCAATACGCTGCAGAAAAGCGCAACCCGCACTTCAGCTTT	1020
Db	3304	ATGCGTCAGTTGTTTCGTCAATACGCTGCAGAAAAGCGCAACCCGCACTTCAGCTTT	3363
Qy	1021	ATCATCAACAGAACGGCATGTTCTCCTTACGTGGCTTGACAAAGAACCAAGTCTCGT	1080
Db	3364	ATCATCAACAGAACGGCATGTTCTCCTTACGTGGCTTGACAAAGAACCAAGTCTCGT	3423
Qy	1081	CTGGCGAAGAGTTTGGCGTATATGCGGTTCTCTTCTGTCGCTGAAATGTGGCCGGATG	1140
Db	3424	CTGGCGAAGAGTTTGGCGTATATGCGGTTCTCTTCTGTCGCTGAAATGTGGCCGGATG	3483
Qy	1141	ACCCAGATTAACATGGGCTCCGCTGTGGGAAGCGATTTGCGCAGTGCCTGTAA	1191

FBI

Db	3484	ACACCAGATAACATGGCTCCGCTGTGGCAAGCGATTGTGGCAGTGCTGTAA	3534
RESULT 6			
	AA546273/C		
ID	AA546273	standard; DNA; 14759 BP.	
XX	XX		
XX	AA546273;		
XX	XX		
DT	18-DEC-2001	(first entry)	
XX	XX		
DE	DNA encoding novel mar regulated protein (NIMR) #42.		
XX	XX		
KW	mar regulated polypeptide; NIMR; microbial infection; antibacteri		
XX	XX		
OS	Escherichia coli.		
XX	XX		
FN	W0200170776-A2.		
XX	XX		
PD	27-SEP-2001.		
XX	XX		
PF	08-MAR-2001; 2001WO-US007478.		
XX	XX		
PR	10-MAR-2000; 2000US-0188362P.		
XX	XX		
PA	(TUFT) TUFTS COLLEGE.		

The invention relates to a method of identifying compounds that modulate a newly identified mar regulated (NIMR) polypeptide activity. The method comprises contacting an NIMR polypeptide with a test compound under interaction conditions, determining the ability of the compound to modulate the activity or expression of the polypeptide, and selecting the modulators. NIMR nucleic acids and polypeptides are used in the treatment of microbial infections, and in screening for modulators of NIMR expression and activity. These modulators can be used to reduce the infectivity of a microbe on a surface, and the virulence of a microbe in a subject suffering from an infection. AAS46232-AAS46278 represent *Escherichia coli* NIMR coding sequences of the invention.

XX
SQ
Sequence 14759 BP: 3703 A: 3840 C: 3713 G: 3503 T: 0 U: 0 Other:

Result No.	Score	Query			DB	ID	Description
		Match	Length	\$			
1	2045	100.0	396	1	AAP71678	Aap71678 aspc gene	
2	2045	100.0	396	2	AW69553	Aw69553 Escherich	
3	2045	100.0	396	2	ADB83285	Adb83285 Escherich	
4	2045	100.0	396	8	ADN18132	Adn18132 Bacterial	
5	2045	100.0	396	9	ADW95404	Adw95404 Amino aci	
6	1931	94.4	402	7	ABO65556	Ab065556 Klebsiell	
7	1682	82.2	397	6	ABM69171	Bm69171 Photorhab	
8	1654	80.9	396	8	ADS42655	Ad42655 Bacterial	
9	1318	64.4	322	8	ADN17578	Adn17578 Bacterial	
10	1317	64.4	423	6	ABP80003	Abp80003 N. gonorr	
11	1311	64.1	397	5	AAU73004	Aau73004 Neisseria	
12	1310	64.1	397	8	ADP08222	Adp08222 Neisseria	
13	1005.5	49.2	394	8	ADS32791	Ad32791 Bacterial	
14	996	48.7	396	8	ADN24844	Adn24844 Bacterial	
15	996	48.7	398	8	ADN22085	Adn22085 Bacterial	
16	982	48.0	397	8	ADN25771	Adn25771 Bacterial	
17	982	48.0	413	7	ABO80253	Ab080253 Pseudomon	
18	975	47.7	397	8	AD326795	Ad326795 Bacterial	
19	975	47.7	397	8	ADS37175	Ad327175 Bacterial	
20	975	47.7	398	8	ADS26427	Ad326427 Bacterial	
21	947	46.3	398	8	ADN24992	Adn24992 Bacterial	
22	924	45.2	395	8	ADN26805	Adn26805 Bacterial	
23	917	44.8	395	8	ADN26569	Adn26569 Bacterial	
24	917	44.8	420	8	ADV07798	Adv07798 Plant fun	

Db 1 MFENITAAPADPILGLADLFRADERPGKINLGIVYKDETKPTVLTSVKKAQYLLENE 60
 QY 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAATAQTPGGTGALRVAADFLAKNTSV 120
 Db 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAATAQTPGGTGALRVAADFLAKNTSV 120
 QY 121 KRWVSNPSPNPKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLFHGC 180
 Db 121 KRWVSNPSPNPKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLFHGC 180
 QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFDFAQGFARGLEDAEGLRAFAAMHKELIV 240
 Db 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFDFAQGFARGLEDAEGLRAFAAMHKELIV 240
 QY 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
 Db 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
 QY 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQEKGNRDFSFIIKQNGMFSFGLTKQVLR 360
 Db 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQEKGNRDFSFIIKQNGMFSFGLTKQVLR 360
 QY 361 LREDFGVYAVASGRVNVAGTMDNMAPLCEAIVAVL 396
 Db 361 LREDFGVYAVASGRVNVAGTMDNMAPLCEAIVAVL 396

RESULT 2

AAW69553
 ID AAW69553 standard; protein; 396 AA.

AC AAW69553;

DT 13-OCT-1998 (first entry)

DE Escherichia coli aspC protein.

XX Brevibacterium lactofermentum; lysC; L-lysine; coryneform bacterium;
 KW aspartokinase; feedback inhibition; dihydrodipicolinate reductase;
 KW diaminopimelate decarboxylase; aspartate aminotransferase.

XX Escherichia coli.

XX EP854189-A2.

XX 22-JUL-1998.

XX 05-DEC-1997; 97EP-00121443.

XX 05-DEC-1996; 96JP-00325659.

XX (AJIN) AJINOMOTO CO INC.

XX Araki M, Sugimoto M, Yoshihara Y, Nakamatsu T;

XX WPI; 1998-379060/33.

DR N-PSDB; AAV40259.

XX Recombinant DNA autonomously replicable in coryneform bacteria - used to
 PT produce L-lysine, codes for e.g. aspartokinase, dihydrodipicolinate
 PT reductase and synthase and di-amino-pimelate decarboxylase.

XX Claim 6; Page 38-39; 59pp; English.

XX The present invention describes a recombinant DNA autonomously replicable
 CC in cells of coryneform bacteria (CB), comprising a DNA sequence coding
 CC for an aspartokinase (AK) in which feedback inhibition by L-lysine and L-
 CC threonine is desensitized, a DNA sequence coding for a
 CC dihydrodipicolinate reductase (DHPRL), a DNA sequence coding for
 CC dihydrodipicolinate synthase (DHPSS), a DNA sequence coding for
 CC diaminopimelate decarboxylase (DAMD) and a DNA sequence coding for
 CC aspartate aminotransferase (AAT). The present sequence represents aspC
 CC from Escherichia coli. The DNA and related products from the present

CC invention, can be used for improving L-lysine productivity by CB. The L-
 CC lysine produced can be used as a fodder additive

SQ Sequence 396 AA;

Query Match 100.0%; Score 2045; DB 2; Length 396;

Best Local Similarity 100.0%; Pred. No. 2e-197;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFENITAAPADPILGLADLFRADERPGKINLGIVYKDETKPTVLTSVKKAQYLLENE 60

Db 1 MFENITAAPADPILGLADLFRADERPGKINLGIVYKDETKPTVLTSVKKAQYLLENE 60

QY 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAATAQTPGGTGALRVAADFLAKNTSV 120

Db 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAATAQTPGGTGALRVAADFLAKNTSV 120

QY 121 KRWVSNPSPNPKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLFHGC 180

Db 121 KRWVSNPSPNPKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLFHGC 180

QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFDFAQGFARGLEDAEGLRAFAAMHKELIV 240

Db 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPDFDFAQGFARGLEDAEGLRAFAAMHKELIV 240

QY 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300

Db 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300

QY 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQEKGNRDFSFIIKQNGMFSFGLTKQVLR 360

Db 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQEKGNRDFSFIIKQNGMFSFGLTKQVLR 360

QY 361 LREDFGVYAVASGRVNVAGTMDNMAPLCEAIVAVL 396

Db 361 LREDFGVYAVASGRVNVAGTMDNMAPLCEAIVAVL 396

RESULT 3

ADB83285

ID ADB83285 standard; protein; 396 AA.

XX ADB83285;

XX 04-DEC-2003 (first entry)

XX Escherichia coli aspartate aminotransferase.

XX aspartate aminotransferase; threonine; fermentation; enzyme.

XX Escherichia coli.

XX WO2003072786-A1.

XX 04-SEP-2003.

XX 25-FEB-2003; 2003WO-JP002067.

XX 27-FEB-2002; 2002RU-00104983.

XX (AJIN) AJINOMOTO CO INC.

XX Akhverdian VZ, Savrasova EA, Kaplan AM, Lobanov AO, Kozlov YI;

XX WPI; 2003-721777/68.

DR N-PSDB; ADB83284.

XX Industrial production of L-threonine by fermentation using Escherichia
 PT modified to enhance aspartate aminotransferase activity, with improved
 PT productivity.

PS Claim 7; Page 20-21; 26pp; Japanese.

XX

This Page Blank (uspto)